THE UNIVERSITY OF

WARWICK

Higher Education in the UK and the Market for Labour: Evidence from the Universities' Statistical Record

by

Luca Mancini

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Economics

University of Warwick, Department of Economics

August 2003

A Margherita, Amelia e Franco, in memoria

To the memory of Margherita, Amelia and Franco

Table of Contents

List of Tables and Figures

Acknowledgments

I wish to acknowledge the ESRC for its financial support and for making this research project possible.
I am particularly grateful to my supervisors Jeremy and Robin for their constant guidance and extensive support throughout my PhD.
I would like to thank the persons whom I met at conferences and, especially, Prof. Ruud, Prof. Shackleton and Prof. Sloane who have made valuable comments to my presentations. I would also like to express my appreciation to Prof. Shafer, Prof. Rumberger, Dr Sianesi, and Prof. Ichino who have offered valuable help by e-mail.
I wish to express my gratitude to colleagues at the Department of Economics and, in particular, to Gaelle, Pedro, Tina and Iwan for listening to my work and providing me with new and interesting perspectives. A special thank you goes to Massimiliano Bratti, for being always generous of comments and ideas, an excellent research partner and a true friend.
I wish to acknowledge the USR, as the original depositors, and the UK Data Archive for the use of the dataset SN: 3456 Universities' Statistical Record. I am also indebted to Abigail McKnight for the use of the data on occupational earnings. None of these individuals or organisations bear any responsibility for the analyses or interpretations presented in the Thesis.
I would like to express my deep appreciation towards Prof. Norman Ireland for being a constant point of reference for me. I owe a special debt of gratitude to Dr Francesco Galassi, for being an excellent supervisor and a very good friend.
I would like to thank Adolfo, Mike, Ivan, Caroline, Miguel, Michela, Cristiano, Giovanni, Silvio, Max, Maru, Ariel, Angie, Claudia, Lucia, Amit, Loles, Tina, Silvia, Toby, Tigran, Rachel, Frederik and Peter for sharing with me so many memorable moments during these years.
I will be eternally grateful to my parents and family for blessing my decision to study abroad, and for their constant encouragement and support.
My last and most special thought is for Hanne, the sweet angel that made this thesis spread its wings.

Declaration

This Thesis is my own work and has not been submitted for a degree at another university. Chapter 5 is closely related to a paper co-authored with Massimiliano Bratti, a fellow PhD student in the Department of Economics of the University of Warwick. Each of us contributed 50% to the paper. My contribution to Chapter 5 is more than 50%, owing to the fact that the analysis on female graduates, and particularly Sections 5.7 and 5.7.1, is entirely my own work.

Abstract

The Thesis seeks to make a contribution to our current understanding of the complex relationship between higher education and the graduate labour market in the UK on both a methodological and policy level. Using administrative data from the Universities' Statistical Record (USR) on complete cohorts of individual students who left university between 1980 and 1993, the Thesis develops along three main avenues: i) identifying the key determinants of graduates' first destinations (Chapters 2 and 3); ii) comparing alternative indicators of employment-related university performance and assessing their robustness to data aggregation (Chapter 4); iii) estimating the differences in graduates' occupational earnings by degree subject (Chapter 5).

The study on first destination considers a broad range of possible outcomes distinguishing between temporary and permanent as well as 'graduate' and 'non-graduate' employment, professional training and postgraduate study, involuntary unemployment and unavailability for work. The analysis reveals significant effects on graduates' employability associated with gender, university type, degree subject, degree class, socio-economic background, and prior qualifications (Chapter 2). Moreover, the impact of all the main factors affecting graduates' early careers has a significant correlation with the business cycle (Chapter 3).

In Chapter 4 we compare employment-related university performance indicators constructed from student-level and university-level data, respectively. Despite student-level data on university statistics now being publicly available, institutions are currently assessed according to indicators based on university-level data, implicitly obtained by averaging over individuals the corresponding student-level information. We find significant differences between the two sets of indicators and argue that the observed discrepancies are the result of an aggregation bias. A Monte Carlo experiment is used to test the validity of this conclusion.

Finally, Chapter 5 looks at the differences of graduates' occupational earnings by degree subject using USR and NES data from 1980 to 1993. We discuss the issue of self-selection of students into the subject of study and apply three alternative modelling strategies to control for self-selection: the proxy and matching method, propensity score matching and a simultaneous equations model accounting for 'selection on unobservables'. The evidence suggests the presence of a significant selection bias originating from the unaccounted correlation between unobservable individual characteristics affecting both occupational earnings and subject choice. Moreover, the ranking of university subjects changes over time.

Glossary of Acronyms and Abbreviations

BTEC	Business and Technician Education Council
CAS	Career Advisory Services
CVCP	Committee of Vice-Chancellors and Principals
Ex-CAT	Ex Colleges of Advanced Technology
FDS	First Destination Survey
GCE	General Curriculum of Education
GHS	General Household Survey
GNVQ	General National Vocational Qualification
HEFCE	Higher Education Funding Council for England
HESA	Higher Education Statistics Agency
HNC	Higher National Certificate
HND	Higher National Diploma
ID.Vs	Identifying variables
KOS	Key List of Occupations for Statistical Purposes
LFS	Labour Force Survey
NCDS	National Child Development Survey
NES	New Earnings Survey
PISG	Performance Indicators Steering Group
QAA	Quality Assurance Agency
RAE	Research Assessment Exercise
SCE	Scottish Certificate of Education
SOC	Standard Occupational Classification
UCAS	Universities and Colleges Admission Services
UCCA	Universities' Central Council on Admissions
UGC	Universities' Grant Committee
USR	Universities' Statistical Record
UUK	Universities UK
YTS	Young Training Scheme

Chapter 1

Introduction

1.1 Introduction

In recent years higher education has been a central issue on the UK political agenda and has stimulated a considerable debate within and between different spheres of stakeholders. From being the preserve of an elite, higher education has evolved into a mass system attracting over 40 percent of young Britons and it is now projected towards becoming increasingly universal.[1] UK higher education has also become more diversified in the composition of the student population owing to the above-average growth of under-represented groups like women, mature, and part-time students, and increasingly global as the result of the greater cross-country circulation of international students and staff.

These structural changes have invariably affected the way the labour market perceives higher education as well as the way the sector itself operates. Achieving a university degree is *per se* becoming much less of a distinguishing mark in the eyes of employers, who are additionally looking at how students performed in their degree, what they studied and which institution they graduated from. Students, for their part, are aware that graduating with a good degree class, as well as choosing 'this' specific course at 'that' particular university will be more decisive for their career prospects than ever before. Universities are increasingly competing against each other to secure the most talented students and the best-trained staff, behaving like enterprises in a 'marketplace' where quality bestows prestige and prestige attracts funding and wealth in a self-reinforcing way. Interestingly, these competitive forces are developing within a system that

[1] The current Education Secretary Mr Charles Clarke has recently reaffirmed in his 2003 White Paper on higher education presented to the Commons on Wednesday January 22, 2003 the government's commitment to increase participation in higher education towards 50% of those aged 18-30 by 2010.

remains highly regulated, where tuition fees and staff pay are decided centrally and taxpayers' money still represents the main source of universities' revenues.[2] Given the resistance by many in the UK to a liberalisation of university fees,[3] the selective allocation of public funding to the sector based on objective indicators of performance like research excellence (RAE) and teaching quality (QAA) currently represents the only institutional response to the emerging tendency towards a greater stratification of universities.

Not surprisingly, these recent trends in higher education have stimulated a wealth of empirical research focusing on the labour market outcomes of university graduates. An overview of the literature suggests the existence of at least two broad areas of research. A first body of studies has been concerned with graduates' employability and encompasses a number of inter-related issues including the determinants and time trends of graduates' first destinations and occupational choices (Rigg *et al.*, 1990; Connor *et al.*, 1997; Purcell and Pitcher, 1996; Pearson *et al.*, 2000; Dolton and Makepeace, 1992), the incidence and the consequences of *overeducation* on graduates' career prospects (Sloane, Battu, and Seaman, 1999; Dolton and Vignoles, 2000), and the construction of employment-related indicators of performance for higher education institutions (Johnes and Taylor, 1990; Smith, Naylor, and McKnight, 2000; HEFCE, 2001 and 2002). A second sizeable strand of the literature has focused instead on estimating the

---

[2] In the academic year 1999-2000 of the £12.8 billion received by universities and colleges in funding, over 60 per cent came from higher education funding bodies and other governmental sources (Source: HESA statistics)

[3] In 1998 flat rate means-tested tuition fees for undergraduate students were introduced in the UK, to ease the mounting financial pressure faced by higher education institutions following the combined effect of the declining level of public spending per student and the explosion of student numbers. The poor results achieved by the reform have led many to ask that universities be set free to decide their fees.

(private) economic return to an undergraduate degree (Blundell *et al.*, 1997 and 2000; Naylor, Smith and McKnight, 2002; Dolton and Makepeace, 1990). The evidence produced by some of these studies has influenced education policy directly. For instance, the decision to introduce tuition fees for undergraduate students in 1998 in the UK partly reflected the evidence, based on empirical estimates, that students receive sizeable returns to a first degree and could therefore afford to make a greater contribution towards meeting the growing financial needs of the higher education sector (Dearing, 1997).

Within this lively research context and fast-changing institutional framework, this PhD Thesis aims to contribute to the existing literature in various ways. Chapter 2 presents an empirical investigation of the determinants of graduates' early career trajectories. Using individualised data on a large sample of individuals who graduated in the UK in 1993, we estimate multinomial logit models of first destinations for all students, for each gender, as well as for specific degree courses. The analysis aims primarily to fill a gap in the literature which, unlike equivalent work on school leavers (Micklewright, 1989; Whitfield and Wilson, 1991; Andrews and Bradley, 1995; Murphy and Shuttleworth, 1997), has either been descriptive in nature or has generally regarded first destination analysis as instrumental to estimating the return to a university degree or to the construction of employment-related performance indicators, rather than a topic worthwhile addressing in its own right. This is surprising if one considers that students are increasingly behaving as forward-looking consumers carefully evaluating the career prospects associated with their educational choices and therefore demanding precise information on how employment prospects differ across universities, subjects, degree classes, socio-economic backgrounds, prior qualifications, age bands and gender groups. At the policy level, the study aims to inform some central issues currently under debate in the UK, like the reform of student finance and the reform of secondary school curricula.

Investigating the determinants of graduates' first destinations is also interesting from a dynamic perspective as the influence of course-related and personal attributes on early career prospects is likely to have changed over time. Chapter 3 extends the first destination analysis to fourteen consecutive cohorts of graduates who left university between 1980 and 1993. Adding a time dimension to the analytical framework means more than simply replicating results for earlier cohorts of university leavers. Supply and demand-side shifts in the graduate labour market can alter the impact of human capital inputs like the university attended, the subject studied, and degree class obtained on graduates' early careers. For instance, during economic slowdowns 'cream-skimming' recruitment selecting individuals on the basis of specific human capital endowments is likely to become more widespread. Likewise, when demand is slack the presence of family or business networks may enhance the influence of socio-economic factors on graduates' early career outcomes (Dolton, O'Neill, and Sweetman, 1996). Broadly speaking, Chapter 3 aims to provide a framework to understand and explain the trends in the UK graduate labour market during the 1980s and early 1990s. The demand and supply of graduates depend on time-specific factors, which are directly related to the business cycle as well as to the changing institutional framework of the higher education system. New entrants to the labour market can be particularly vulnerable to demand and supply shifts and starting a career in an adverse economic climate can prove particularly difficult because graduate recruitment is often one of the first casualties of a downswing in demand. Similarly, supply-side shocks like a surge in enrolment rates or changes in the private cost of a university degree can lead to short-term surpluses or shortages of highly educated labour. Finally, graduates may become increasingly unable to secure occupations that command the pecuniary return they expect from a higher education qualification. The explosion in participation has generated widespread concern that the economy is producing university graduates at a faster rate than it is creating suitable jobs for them. Arguably, recently coined concepts such as 'overqualification' or 'overeducation' are best understood and appreciated in a dynamic context.

In Chapter 4 we contribute to the literature on performance indicators by considering the effects of data aggregation on the construction of employment-related measures of university output. One fundamental difference between two of the most influential UK studies on employment-related university performance indicators authored by Johnes and Taylor (1990) and Smith *et al.* (2000) respectively, is that the first uses information at the university level, while the latter is based on individualised student data. Despite the fact that 'micro' level information on university students is now publicly accessible, institutional performance continues to be widely assessed on the basis of 'macro' or university-level data. The latter is implicitly obtained in a 'representative agent' fashion by averaging over individuals the corresponding student-level information. There is a sizeable theoretical literature (Stoker, 1982; Van Daal and Merkies, 1984; Lewbel, 1992; Deaton and Muellbauer, 1980; Richards and Ben-Akiva, 1975) warning of the pitfalls of aggregating individual decisions to predict macro behaviour. Empirical research on transportation mode choices has also shown that the estimation of macro behaviour is highly sensitive to the way individual characteristics and decisions are aggregated (Talvitie, 1973; Koppelman, 1976; Westin, 1974; Watson and Westin, 1975; Nam, 1997). Given that universities are understandably sensitive to their position in the tables, an analysis of whether performance indicators are robust to data aggregation appears warranted. In fact, the publication of league tables invariably creates winners and losers and the release of flawed information in an increasingly competitive but still regulated 'marketplace' such as UK higher education can have scarring consequences on the sector. Using student-level data on 1993 university graduates we compare employment-related university performance indicators constructed from micro and macro data, respectively. A Monte Carlo experiment is used to test whether the differences between the two sets of indicators are the result of an aggregation bias. We also assess the impact that alternative aggregation techniques that consider higher moments than the mean (Taylor series expansions) or procedures that partition the student population into more homogeneous classes (classification) have on the aggregation bias.

Chapter 5 contributes to the return-to-education literature by estimating relative earnings *premia* by subject studied over the period 1980-1993 correcting for the endogeneity of subject choice. In a recent review of related empirical studies for the UK, Chevalier *et al.* (2002) highlight how the rate of return to a university degree is likely to differ substantially across academic fields. Although these studies recognize that subject choice may be endogenous to earnings determination, they make no direct attempt to address the issue of student self-selection into academic subjects and its consequences for the estimation of the rate of return to specific degree courses. Blundell *et al.* (2000) using National Child Development Study data (NCDS) assume that the wealth of information contained in the dataset is sufficient to correct for the endogeneity of subject choice by 'matching' individuals from different subjects on a comprehensive set of observable factors adequately 'proxied' in the data (*matching and proxying* method). By relaxing this assumption, we take a more general approach towards the problem of the potential endogeneity of subject choice. In particular, the earnings *premia* estimated by the *matching and proxying* method are contrasted with estimates obtained from two alternative approaches: i) 'propensity score matching' methods where the individual's subject studied and average occupational earnings represents the 'treatment' and the 'outcome', respectively and ii) simultaneous equations models of subject choice and earnings determination. Moreover, the dynamic dimension of the analysis is expected to complement the study on graduates' employability discussed in Chapter 3, and contribute towards a better understanding of the trends in the graduate labour market in the UK between 1980 and 1993.

Finally, Chapter 6 concludes by summarising the main results and indicating possible extensions and directions for future research.

1.2 Data

Given the uniqueness of the data used and their centrality to all the analyses presented in this Thesis, this section is dedicated to illustrate the main features of the Universities' Statistical Record (USR) dataset. The choice to include the description of the data at this early stage also reflects an attempt to enhance the fluency of the exposition by avoiding redundancies and repetitions in the next chapters. The research is conducted using individualised administrative data from the USR on undergraduate students who graduated in the UK between 1980 and 1993. The Record was set up in 1968 under the joint auspices of the University Grant Committee (UGC) and the Committee of Vice Chancellors and Principals (CVCP). The data are aggregated in three volumes. Volume 1 contains information on staff and students (both undergraduate and postgraduate), Volume 2 holds information on the first destinations of university leavers, while Volume 3 contains financial information on universities. The USR was the depository of all university statistics from the academic year 1972/73 until 1993/94, when the Record was taken over by the Higher Education Statistics Agency (HESA).[4] The Record is undoubtedly the most comprehensive source of information on individual university students in the UK and it is still largely under-utilised by

[4] Owing to major changes after 1980 in the USR coding procedure of some key first destination variables used in this study, we use only data starting from 1980 to ensure inter-cohort comparability.

researchers and practitioners, owing to the relatively recent disclosure of its archived files.

The information on undergraduate students can be grouped into four categories:

i. *Personal information*: date of birth, gender, marital status, country/county of domicile, country of birth, occupation of parent or guardian;

ii. *Academic history*: last full-time school attended, other full-time/part-time post secondary educational institution attended, General Curriculum of Education (GCE) A-level or Scottish Certificate of Education (SCE) Higher grade results, other entry qualifications like General National Vocational Qualifications (GNVQ) and Business and Technician Education Council (BTEC), course for which admitted;

iii. *Annual information*: university attended, subject of course, legal duration of course, year of course, date of enrolment, method of study (full-time, part-time, sandwich), qualification aimed for;

iv. *Leavers' details*: qualification obtained, class of degree, date of leaving, reason for leaving, first destination information (including primary classification, type of occupation, type of employer).

First destination information is based on self-reported answers to a postal questionnaire sent by the Career Advisory Services (CAS) of each university to their newly qualified graduates approximately six months after graduation, and requesting details about their current labour market status. The First Destination Survey (FDS) is conducted with a common format by each CAS and the results are processed centrally[5]. The response rate to the FDS is about 80% of the eligible population of all university leavers and has remained remarkably stable over the 1980s with evidence of a decline after 1991. Despite a less than complete response rate, the USR do give detailed information on a much larger sample of university graduates than is available elsewhere. The limitations of the dataset associated to survey non-response and to the use of first destination information collected only six month after graduation are carefully discussed in Chapter 2.

A particularly useful feature of the USR data is its 'mergeability', i.e. the possibility to merge complementary information into the Record from other sources. For instance, crucial information on graduates' occupational earnings and parental social class is not found in the original USR files. However, we were able to merge these variables into our dataset by matching the type of occupation classification of the student and her or his parents with the Standard Occupational Classification (SOC) used in the New Earnings Survey (NES). The details on the mapping between USR and SOC occupational classifications are given in Chapter 5 (Section 5.4).

Finally, a major effort has been made to ensure comparability of degree subjects during the period 1980-1993.[6] After the classification of academic subjects was overhauled in 1985, it was necessary to bridge the old course codes with the new classification system. This bridging exercise have been made possible, on the one hand, by the availability in the USR of highly disaggregated old and new subject

[5] In general, for each of the three Volumes data collection is conducted according to standard procedures, which include an automatic update process, computer validation, error listings, feedback tables and a final certification of quality by the USR Correspondent from each university. This considerably improves consistency across universities and reduces attrition problems over time.

codes (4-digit level), and by the availability of annual information detailing for each individual the subject studied from enrolment to graduation, on the other hand. Therefore, for individuals who enrolled before 1985 and graduated after 1985 we know both the old and new subject classifications. This has helped make an accurate conversion from old-to-new codes for most of the main subjects and avoid discontinuity between cohorts.[7]

[6] Consistency over time has also been a key issue with respect to the coding of the occupational earnings series (see Section 5.4 for details).

[7] The new classification was characterised by a more diversified range of courses, especially with respect to interdisciplinary or combined subjects. Those codes for which attrition problems were encountered were re-coded as 'Other' and should be considered as residual categories.

Chapter 2

After university: an empirical investigation on graduates' first destinations in the UK

2.1 Introduction

This chapter investigates the determinants of university graduates' first destinations in the UK. Our primary aim is to ascertain which personal attributes and course-related characteristics make students more likely to enter the labour market or, alternatively, which factors make students more likely to defer entry to a later date, either because they continue their studies or because they take longer to make the transition from university to work. Compared to the literature on the first destination choices of UK school leavers (Micklewright, 1989; Whitfield and Wilson, 1991; Andrews and Bradley, 1997; Murphy and Shuttleworth, 1997), university leavers have received less attention. Studies on graduates' first destinations tend to consist either of descriptive reports commissioned by education departments or Career Advisory Services (Rigg *et al.*, 1990; Purcell and Pitcher, 1996; Connor *et al.*, 1997; Pearson *et al.*, 2000), or to focus on the occupational choices of those already employed (Dolton, Makepeace, and Van der Klaauw, 1989; Dolton and Makepeace, 1992), or to correct estimates of the private return to a university degree for the non-random selection into employment (Blundell *et al.,* 2000), or else to construct employment-related performance indicators for higher education institutions (Smith *et al.*, 2000; McKnight, 1999).

We believe that a more comprehensive study focusing on the determinants of graduates' first destination outcomes is interesting in its own right. The increasing cost of higher education borne by students in the UK and the greater competition and fragmentation in the labour market for the highly qualified is expected to make individuals more selective in their choices and demand more precise information on how their career options after graduation are going to be affected by factors like prior qualifications, degree performance, subject studied or the institution attended.

More precisely, the chapter aims to contribute to the existing literature in at least three directions. First, we extend to university leavers the approach used by Andrews and Bradley (1997) for school leavers, which consists of modelling a wider spectrum of possible destinations open to newly qualified graduates than usually found in previous studies.[1] We distinguish between permanent and temporary employment, professional training and postgraduate studies,[2] unemployment and non-employment (in the sense of unavailability to work, study, or training). Each of the above six destinations is likely to reflect a unique mix of personal and course characteristics as well as different individual discount rates and preferences.

Second, borrowing from a methodology widely used in the literature to decompose the gender wage gap first proposed by the seminal articles of Oaxaca (1973) and Blinder (1973), the chapter aims to ascertain to what extent male-female differences in first destination probabilities are explained by differences in observable characteristics, rather than being the result of unobservable factors such as potential labour market discrimination.

Third, we use a finer definition of academic subjects than usually found in related studies. Despite the wide consensus on the influential role of subject studied on graduates' employability (Dolton, Greenaway, and Vignoles, 1997), sample size limitations have often forced researchers to consider broadly defined academic

[1] The set of alternative outcomes is usually restricted to employment, further study or training and unemployment (Johnes and Taylor, 1990; McKnight, 1999).

disciplines which can be highly heterogeneous with respect to first destination outcomes (Blundell *et al.*, 1997 and 2000; Johnes and Taylor, 1990; Dolton *et al.*, 1989). We are able to overcome these limitations by using a large individualised dataset containing detailed course information on a complete cohort of students who graduated in the UK in 1993. The size and quality of the dataset has afforded us to consider some popular degree courses at the departmental level (2-digit level of the UCAS classification) to account for potential differences in first destination behaviour among graduates from different departments within the same faculty.

A subject-specific analysis is particularly well suited to inform some of the current debates on education policy in the UK, like the reform of student finance and the reform of secondary school curricula. With respect to the first issue, the proposal to set universities free to offer different courses at different prices depending on demand has raised widespread concern in the UK. Opponents of the reform argue that *top-up* fees may adversely affect participation and educational choices of individuals from less affluent social backgrounds. In this chapter we aim to establish whether this concern is justified. With regard to the second issue, critics claim that the English secondary school system, and in particular the curriculum at A-level, inadequately prepares students for the world of work. One recurrent argument is that the curriculum for 16-19 year olds is too narrow and provides overly specialised knowledge. If individuals who take a narrower curriculum at 16-19 also tend to take certain degree courses, conditioning on degree subject will be important to ascertain the effect of curriculum breadth on graduates' first destinations.

---

2 Professional training encompasses all further study routes other than for a degree qualification,

The rest of the chapter is structured as follows. Section 2.2 reviews previous UK studies on the first destinations of school/university leavers. The modelling strategy is illustrated in Section 2.3. Section 2.4 presents the dataset and the variables used in the analysis, while Section 2.5 describes the sample and the main features of the data. Section 2.6 discusses the estimation results for the whole sample of graduates, while Section 2.7 presents an in-depth analysis on gender effects and differences. After breaking down the sample by degree course, Section 2.8 takes a closer look at the effect of subject studied and provides in turn some insights into the reform of student finance (2.8.1) and the reform of A-level curriculum (2.8.2). Finally, Section 2.9 concludes with a summary of the main findings and sets up the directions for the next chapter.

2.2 Literature review

As anticipated in the introduction, much of the existing literature on the transition between education and work in the UK has focused on school leavers. Given its relevance to the modelling strategy used in this chapter, Section 2.2.1 will review some of these studies separately from research focusing on university leavers, which will be dealt with in Section 2.2.2.

---

both in the UK and overseas. On the other hand, postgraduate study includes primary degree and higher degree study/research, both in the UK and overseas.

2.2.1 School to work transition

The empirical literature on the school to work transition in the UK has a long-standing tradition. A number of studies have concentrated on the staying on decision of school leavers. Pissarides (1981) using aggregated time series data for England and Wales between 1955 and 1978 published by the Department of Education and Science found that staying on rates tend to be counter-cyclical reflecting a dominant discouraged worker effect.

Micklewright (1989) estimated reduced-form logit models of the probability of leaving school at sixteen in England and Wales using micro data from the National Child Development Study and found that the decision to stay on is crucially affected by socio-economic factors like parental education and social class, even after controlling for children's academic ability and type of school attended.

Whitfield and Wilson (1991) extended Pissarides' work by accounting for the effect that the introduction of the Young Training Scheme (YTS) in the 1980s had on 16-year-old's staying on decisions. They found that the YTS had significantly reduced numbers remaining in full-time education.

Andrews and Bradley (1997) criticise the empirical approach used in these studies where staying on decisions are modelled as a simple binary choice and youth training and unemployment are typically entered as regressors rather than being regarded as alternative destination outcomes to be jointly determined by the model. In their study, the transition from school is modelled as a choice between six mutually exclusive and exhaustive outcomes including two continuing education options (non-vocational and vocational), youth training, two employment options (with on-the-job training or general-skills training), and unemployment. Using Career Service data on Lancashire school leavers in 1991, they estimate multinomial logit models and find that the probability to stay on is significantly affected by academic performance, school-specific factors like type and size, and expected lifetime earnings. Their findings also suggest that the outcomes can be ordered by decreasing levels of student ability, leading to the conclusion that the youth labour market is segmented.

Murphy and Shuttleworth (1997) estimate nested logit models of first destination using cross-section survey data on 1991 school leavers in Northern Ireland. Six outcomes are considered: employment, youth training, unemployment, higher education, further education, and other inactivity. Amongst other results, the authors find that religion is a key factor determining first destinations, even after controlling for family background and academic ability.

2.2.2 Graduates' first destinations

Compared with the established UK literature on the school-to-work transition, its counterpart focusing on the transition from university to work does not lend itself to be confined into a homogeneous body. A sizeable strand of this literature consists of descriptive reports usually based on *ad hoc* first destination surveys and it is primarily aimed at comparing the time trends in the demand and supply of different groups of university graduates as identified by their occupation and type of employer, academic subject, gender, age, and social background (Brennan *et al.*, 1993; Connor *et al.*, 1997; Pearson *et al.*, 2000).

Elias (1999) explores the extent and the determinants of unemployment in the early careers of newly qualified leavers from UK higher education institutions in 1995. Using survey data from a selected pool of UK universities, he finds that certain personal characteristics like being male, over 47 years of age at the time of graduation, of 'non-white' ethnic origins, or coming from parents who were not in employment, tend to correlate with a worse than average unemployment record up to 3.5 years after graduation. With respect to the influence of course-related characteristics, the study shows that individuals who graduated from vocational courses like Education and Medicine were significantly less likely to be unemployed. Similarly, individuals graduating with first class honours are more than twice as unlikely to experience joblessness compared to people with a lower second or a third.

Purcell and Pitcher (1996) investigate the early labour market experiences of 1996 UK-domiciled graduates 1.5 years after graduation. Using survey data on individuals from 21 institutions, the study builds on an earlier survey reporting the career expectations of the same individuals while they were in their final undergraduate year. Results show that the subject studied at university is the single most important determinant of the ease or difficulties to start a career. A degree in Engineering, Computer Science or Business Studies is associated with positive career outcomes and high levels of job satisfaction. By contrast, graduates in Humanities, Social Sciences and Life Sciences tend to experience more difficulties starting a career and are also more likely to undertake postgraduate education. Other findings indicate that although a good degree class is essential to enter a 'graduate-level' job, a high proportion of graduates with lower second class honours or less reported to be in occupations which they regarded as appropriate to their degree qualification.[3] Male graduates are more likely than females to be in full-time employment, while mature students take longer to establish a long-term career.

Although this literature provides useful insights into the heterogeneity of students' early careers stemming from graduates' differences in personal and course-related characteristics, relatively few studies have used logistic regression analysis to estimate the impact that individual attributes exert on graduates' first destination choices. McKnight (1999), using the same dataset as in Elias (1999), estimates logit models of the probability of entering a 'non-graduate' occupation against a wide range of individual-specific attributes. The results show that the odds of finding employment in a 'non-graduate' occupation are approximately 50% higher if the graduate is female, and 36% higher if the individual received on average less than 24 points in her best three A-level passes.

Smith *et al.* (2000) use student-level data on 1993 university leavers to estimate a binomial probit model of the probability that the university graduate is either employed or in further study rather than unemployed or inactive six months after graduation. Their results reveal significant effects associated with age, A-level curriculum and performance, social class, degree class and subject of degree.

Dolton and Makepeace (1992) using a survey of 1980 graduates estimate a multinomial logistic model of the probability of entering one of six alternative occupational categories (teachers, scientists & technicians, managers, public

[3] Occupations for which graduates are 'overqualified'. For a detailed definition see Chapter 3, Section 3.7.1.

administration, self-regulating professions, other). They find significant effects associated with social class and degree class.

2.3 Methodology

For a newly qualified graduate, the choice of what to do after university is arguably the climax as well as the *leit motif* of a sequence of decisions taken earlier in his or her life: the choice to go to university, the choice of schooling track and curriculum at 16-19, and finally the choice of the academic subject and institution to attend.[4] Conditional on his or her past education decisions,[5] each graduating student is basically confronted with a two-way choice: to seek employment immediately or to defer entry into the labour market to a later date. In the latter case, an individual can either decide to invest in more education or training, or to simply pause for a while and take some time out to think about the future. Among those students actively seeking work, not everyone is expected to find a permanent job by the time she or he graduates or soon after. Some graduates may take up temporary posts to pay off their debts and buy some more time to think about longer-term career plans. For others, job search can last for several months after the date of graduation. Therefore, a new graduate can find himself or herself in one of six possible destinations. These are: 1) *permanent employment*, 2) *temporary employment* (work expected to last less than 3 months),

4 Given the increasing competition among applicants for a place at university, particularly in highly ranked institutions, the awarding university may not necessarily be the first choice for some applicants. This may also be true for the choice of the subject, although to a lesser extent.

5 We only consider individuals who went to university and obtained a degree. Therefore, all the results should be considered as conditional to having graduated from university.

3) *vocational training*, 4) *postgraduate study*, 5) *unemployment* (and actively seeking work), and 6) *not seeking work or study/training.*

As Andrews and Bradley (1997) did for school leavers, we first attempt to model all the possible destinations open to a new graduate after university. A six-way logit model is used to determine who are the individuals more likely to fall in any of the first destination outcomes described above. Although some of the above pairs of destinations may be closer substitutes, it is not obvious how one can aggregate alternatives *a priori* without imposing arbitrary restrictions on first destination behaviour. Each destination is likely to reflect a unique mix of personal and course characteristics and different discount rates and preferences. Likelihood ratio tests for equality constraints across different destination pairs are shown in Table 2.1. The tests massively reject any pooling of alternatives, suggesting that the six groups should be jointly considered in modelling the career decisions of university graduates.

However, in spite of this evidence, in Section 2.7 we will aggregate pairs of the above destinations into three main alternatives: 1) *employment* (including permanent and temporary work), 2) *further study or training*, and 3) *unemployment or out of the labour force* (including voluntary and involuntary unemployment). The reason why we estimate this restricted specification is primarily to avoid overly-small cell sizes and the consequent loss of precision in the parameter estimates when we break down the analysis by subject studied. Furthermore, results will be more directly comparable with previous studies on graduates' first destinations (Johnes and Taylor, 1990; McKnight, 1999), which use a similar three-way modelling strategy.

Each graduate faces the following reduced-form choice problem,

$$\begin{aligned} V_{ij} &= f(X_i) + \varepsilon_{ij} \\ p_{ij} &= P(y_i = j) = P(V_{ij} > V_{ik}) \end{aligned} \qquad j = 0,1,\ldots,J; i = 1,2,\ldots,N \qquad (2.1)$$

where y_i is the latent indicator variable and V_{ij} is the value of outcome j to individual i which depends on a set of individual characteristics X_i. The model assumes that any destination outcome has a positive selection probability for each individual. The error term ε_{ij} follows an extreme value distribution and it is independent across alternatives (McFadden, 1978).[6] The model assumes that each individual self-selects into the destination that maximises her expected value or utility conditional on her endowments of productivity-related characteristics X_i. The probability that individual i with characteristics X_i chooses alternative j can be expressed as

$$P_{ij} = \frac{\exp(X_i'\beta_j)}{\sum_{k=0}^{J} \exp(X_i'\beta_k)} \qquad (2.2)$$

To calculate the marginal variation in predicted probabilities induced by changes in X_i, the maximum likelihood estimated coefficients β_j need to be transformed into marginal effects, as

6 We explicitly tested the IIA assumption for both the unrestricted and restricted model using the method suggested by Hausman and McFadden (1984). We found no evidence against the null hypothesis of independence.

$$P_{ij}\left(\beta_j - \sum_{k=0}^{J} P_{ik}\beta_k\right) \quad (2.3)$$

If X_{1i} is a dummy as for most of the variables used in our analysis, the marginal effect measures the difference in the predicted probabilities when X_{1i} varies between 0 and 1, that is

$$P(j=1)_{X_{1i}=1,\ X_{ri}=\tilde{X}} - P(j=1)_{X_{1i}=0,\ X_{ri}=\tilde{X}} \quad i=1,\dots,n;\ j=1,\dots,m \quad (2.4)$$

where the vector of the remaining r-1 covariates X_{ri} takes values $\tilde{X}$, which can either be the vector of the mean values of X_{ri} (average individual) or, more generally, a vector of non-average values identifying a default individual with some specific characteristics. Because the coefficients need not have the same sign as the derived marginal effects, the results are presented in terms of marginal effects rather than estimated coefficients. For the sake of completeness, the coefficients (with standard errors) are reported in Appendix 2B, but they are not discussed.[7]

We are aware of the limitations of the methodology described above. First, we have no information on people who did not go to university. Therefore, as noted earlier, our estimates are to be regarded as conditional on attendance at a university. Additionally, only individuals who graduate take part in the FDS,

[7] Alternative discrete choice models could have been chosen. However, an ordered logistic specification was ruled out because there is no reason to believe that individuals rank first destination outcomes in the same way. Conditional or nested logit models would have been preferred, but could not be implemented due to the unavailability of choice-dependent attributes.

implying that first destination information is unavailable for those students who failed their degree or dropped out of university. This renders our analysis also conditional on graduation.[8]

Second, first destination outcomes refer to arrangements that graduates made by the end of the calendar year in which they gained their degree. Career decisions made within less than six months from graduation may not necessarily reflect future career choices. However, despite this short-term horizon, we believe that first destination outcomes are still a powerful instrument to elicit longer-term career preferences and prospects. Dolton and Makepeace (1992) find that there is little career mobility between graduates' occupations six months and six years after leaving university. Purcell and Pitcher (1996) report that nearly 60% of students had already started seeking employment by the last term of their final year, and more than half planned to embark upon career-related jobs. Moreover, as a more recent study by McKnight (1999) suggests, unemployment 6 months after graduation is a surprisingly good predictor of longer-term difficulties in the labour market.

Finally, we assume that conditioning on the observable attributes X_i is sufficient to control for the potential endogeneity of some choice variables, primarily subject studied. The model relies on the assumption that individuals who are identical in their observable characteristics X_i but choose different subjects do not differ on average in the unobserved dimension ε_i. Clearly, the success of this methodology rests crucially on whether the unobserved factors are adequately

[8] About 4.4% of the 1993 leavers failed their degree, and a further 9%. dropped out of university before completion. Summary statistics reveal that above-average proportions of these individuals are found among

proxied in the data (Blundell *et al.*, 1997).[9] We believe that the wealth and quality of information contained in the USR is probably unmatched in this respect. The next section briefly presents the data and describes the variables used in the analysis.

2.4 The data

With the exception of the study by Smith *et al.* (2000), the empirical literature on graduates' first destination decisions for the UK is either based on data aggregated at the university level (Taylor and Johnes, 1990), or on surveys either targeted at students from narrowly-defined geographical areas (McKnight, 1999) or based on relatively small samples of individuals randomly selected from the graduate population (Dolton and Makepeace, 1992). Our investigation is conducted using individual-level data on the most recent cohort of university graduates available from the USR. Using the wealth of information contained in the Record, we consider the following variables:[10]

i. *Degree course*. Depending on the number of observations and on the policy relevance of some courses, two different levels of aggregation are used corresponding to the one-digit (H for Engineering, for example) and the two-digit (L1 for Economics, for example) UCAS codes;[11]

---

mature students, working class students, Law students, 'Ex CAT' students and individuals without A-level qualifications or with poor A-level performance.

[9] In Chapter 5 we model explicitly the endogeneity of subject choice when estimating relative occupational earnings *premia* by course of degree.

[10] A complete list of the variables used and their definition can be found in the Appendix 2A.

[11] Students who read joint or combined degrees are allocated to the subject to which their primary course refers.

ii. *Institution attended.* The 57 'pre-1992' universities (these exclude the former Polytechnics) listed in the Record have been aggregated into the seven 'types', which identify groups of universities by age of foundation or location;[12]

iii. *Degree class*;

iv. *Social class.* This variable is based on parental occupation obtained by matching USR parental occupation information with Standard Occupational Classification (SOC) from the New Earnings Survey (NES);[13]

v. *Schooling background.* We include controls for GCE A-level and SCE Higher count, subject and grade score, other (than A-level/Higher) entry qualifications like BTEC (Business and Technician National Diploma) HNDs (Higher National Diplomas) or HNCs (Higher National Certificates), and the type of school attended;

vi. *Other course characteristics.* These include controls for the method of study (full-time/part-time) and course duration;

vii. *Personal characteristics.* These include dummies for gender, age, marital status and region of residence prior to university;

viii. *Departmental information.* The subject dummies are interacted with the individual university dummies to generate department-level variables like staff/student ratio, the proportion of postgraduates, the proportion of 'good' students (who graduated with an upper second class or higher), the proportion of males, and the points scored by the department in the last Research Assessment Exercise (RAE). These factors should reflect qualitative differences across

[12] University types rather than individual institutions are chosen because in this chapter we concentrate on the effect of other factors like subject studied and schooling curriculum. However, in Chapter 4 we focus on the effects of graduating from a specific institution. These effects are then used to construct university performance indicators and league tables.

university departments, and are expected to capture other dimensions of the departmental 'value-added', which remain unexplained after controlling for university type and subject studied.

The dependent variable is a six-way categorical variable constructed on the basis of self-reported answers to the FDS. Graduates were classified into one of six mutually exclusive categories corresponding to the groups discussed in the methodology section.

Unfortunately, every year approximately 10% of the graduates who received the first destination questionnaire did not send it back to their CAS.[14] We are aware that ignoring non-respondents can lead to sample selection problems if these individuals represent a 'non-random' section of the graduate population. This issue will be dealt with in the next section, where we also present the salient features of the sample.

2.5 Sample and summary statistics

Our sample consists of 63,706 students who graduated in the UK in 1993 and responded to the first destination survey. The 1993 cohort of university leavers is the most recent wave of graduates on which we have information.[15] Based on their answers to the first destination questionnaire, 28,470 graduates (44.7%) were

[13] We wish to thank Abigail McKnight for making this variable available to us.

[14] Not all university students participate in the FDS. Individuals who fail to graduate (including drop-outs) and most overseas graduates returning to their country of origin do not receive the questionnaire.

[15] In fact, after the USR was taken over by the HESA in the academic year 1993/1994, individual-level data are no longer accessible.

in permanent employment (PWORK) at the time of the survey, 3,786 (5.9%) reported to be in temporary employment (TWORK), 10,219 (16%) engaged in vocational training (TRAIN), 10,750 (17%) went on to study for higher degree qualifications (PSTUDY), 7,298 (11.5%) reported to be still unemployed (UN), and 3,183 (5%) were not actively seeking work or training/postgraduate study (OLF). With respect to social class, graduates coming from the professional (SC I) and intermediate (SC II) parental backgrounds represent over 60% of the entire sample, compared to 17% coming from skilled manual (SC IIIM) and partly skilled/unskilled backgrounds (SC IV-V). Over 80% of the students obtain second class honours either upper or lower (DUSEC and DLSEC), while nearly 60% receive at least an upper second class degree (DFIRST and DUSEC). Graduates from 'old civic' universities (OCIV) constitute the largest group (nearly 35%), while 7% are from 'Oxbridge' (OXBR). Approximately 23% of the students went to an independent school (INDEP). Males outnumber females (54% and 46% respectively). Finally, 11% of graduates were aged 21 or older when they enrolled (MATURE), and less than 2% studied part-time (PTIME).

The sample does not include 6,900 individuals who qualified to receive but did not respond to the FDS. Table 2.2 shows that non-response rates are particularly high among males, mature students, graduates with a poor degree class, and individuals without A-level Mathematics (ALMATH). There is also some evidence of a high proportion of non-respondents among Classics, Humanities and Sociology graduates.[16] Although in these sections of the graduate population there is a relatively high concentration of unemployed and temporarily employed, it is hazardous to assume a systematic correlation between non-response and these two categories.[17] The sample also excludes medical and non-UK graduates.[18] Medical students were not included because their academic profile and first destination behaviour are very different from other courses. Medicine graduates are generally awarded a 'pass' degree class and almost all enter an occupation after graduation.[19] Non-UK students were excluded because first destination information is available only for the minority of individuals that remain in the UK, which is not expected to be representative of the whole category.[20]

2.6 Estimation results for all students

Table 2.3 shows a selected set of marginal effects for the entire sample of 1993 graduates calculated at the sample means of X_i.[21] For the sake of completeness, the full set of estimated coefficients (with standard errors and diagnostics) are

---

[16] The incidence of non-response is expected to be significantly mitigated by aggregating individual universities into institution types, due to the potential differences between institutions in the zeal with which they pursue non-respondents.

[17] A test (Hausman and McFadden, 1984) of the hypothesis of irrelevance of non-response as an additional category for the dependent variable was not rejected, and therefore the exclusion of these individuals from the analysis is not expected to undermine the main findings.

[18] By medical degrees we strictly refer to Medicine and Dentistry courses which are classified under the macro group 'A' of the one-digit UCAS classification of academic subjects. This means that graduates from courses allied to medicine (one-digit UCAS code 'B') like Pharmacy or Physiology are retained in the sample as well as graduates in agriculture-related subjects (one-digit UCAS code 'D') like Veterinary Science.

[19] In the UK a medical degree takes usually 5 year to complete (2 'pre-clinical' years plus 3 'clinical' years). At the end of their second year, successful students are given the opportunity to take an extra year to complete a first degree in science, before resuming their medical degree.

[20] This happens because Career Offices tend to lose track of those graduates who return to their countries of origin.

[21] The marginal effects are calculated using Equation (2.3) for continuous variables and Equation (2.4) for categorical variables.

reported in Appendix 2B, but they are not discussed.[22] The main results are summarised below:

i. *Subject studied.* Relative to Economics (default course), graduates in Education (EDU), Allied Medicine (ALMED), Computer Science (COMP), Business (BUS), and Engineering (ENGIN) are significantly more likely to enter permanent employment. For instance, a Business graduate is, on average, 8 percentage points more likely to enter permanent employment than an otherwise 'observationally identical' Economics graduate. For all other courses we find significant negative effects. This result is in line with the findings of previous studies (McKnight, 1999) and reflects the vocational nature of these courses, which reduces the range of occupational opportunities, on one hand, but makes the graduate more readily employable, on the other. Both these factors are expected to shorten the duration of the job search. The probability of temporary employment is relatively high for graduates in Sociology (SOCIO) and Politics (POL). Not surprisingly, Law graduates are 58 percentage points more likely than Economists to enter professional training. Large positive effects are also found for Modern European languages (MEUL), Classics (CLAS) and Humanities (HUM). USR information on graduates' type of work reveals that legal (solicitors and barristers) and teaching (primary and secondary schools) training attract over 70% of all graduates engaged in training programmes. This explains why training is so popular a destination for graduates in these courses. Graduates in science-based degrees, and in particular Chemistry (CHEM), Biology (BIOL) and Physics (PHYS) are up to 30 percentage points more likely to undertake postgraduate

[22] The statistical significance (t-statistic) of the estimated coefficients is also assumed for the corresponding marginal effects.

education compared to the otherwise 'observationally identical' Economics graduate. These findings may reflect the fact that science-based departments are involved, more than other departments, in high-budget research projects which translate into a greater amount of funding available to postgraduate research. Therefore, newly qualified science graduates are more likely to be offered competitive research contracts, which may entice them to enrol in postgraduate programmes;

ii. *University type*. Graduates from 'Oxbridge' (OXBR) or Ex Colleges of Advanced Technology (EXCAT) institutions are significantly more likely to find a job relative to the 'observationally identical' individual graduating from 'old civic' (OCIV) institutions (default), even after controlling for subject studied and degree class. Screening effects associated with the reputation of 'Oxbridge' colleges on one hand, and the links to the surrounding industrial community for EXCAT institutions on the other hand, are likely to be important factors behind these results;

iii. *Degree class*. Not surprisingly, degree performance has a positive and monotonic effect on the probability of moving on to higher degree courses. In particular, a graduate with a first class degree (DFIRST) is 15 percentage points more likely to become a postgraduate vis-à-vis an otherwise 'observationally identical' individual with upper second honours (DUSEC) – the default case. It is also evident that the lower the degree class the higher is the probability of being unemployed or employed in a temporary job. Interestingly, degree performance has a weak effect on the probability of permanent employment. This can be due to the large and monotonic effects associated with A-level performance (SCOREA-

SCORED). The fact that A-level grades are an important predictor of future university performance (Smith and Naylor, 2001) partly explains the weak effects of degree class once performance at A-level is accounted for. It is also possible that imposing equality restrictions across subjects and gender conceals important degree class effects. For this reason, in the next sections these constraints will be removed by running separate regressions for each gender and degree subject;[23]

iv. *Social class*. Students with professional parents (SC I) are more likely to undertake further studies relative to the default category (SC II). On the other hand, graduates from unskilled or partly skilled parental backgrounds (SC IV-V) are 3.3 percentage points more likely to be unemployed. Again, imposing equality constraints across subject studied may cloud more significant social class effects. Given the policy relevance of these issues, in the following section we will explore the effect of social class in greater detail;

v. *Gender and other personal characteristics*. Female graduates are significantly more likely to enter employment and training courses and less likely to be unemployed or in postgraduate education than males, even when controlling for differences in subject mix and degree class. Interestingly, marital status (MARRIED) has a strong positive effect on the probability of permanent employment, but only for males. In fact, when marital status is interacted with gender, married female graduates are nearly 14 percentage points less likely to enter a permanent job than their single counterparts and 11 percentage points

---

23 We acknowledge the potential endogeneity of degree performance with respect to first destination decisions. An obvious source of selection bias originates from unobserved heterogeneity in effort levels. However, the inclusion of factors such as A-level scores and count, social class, and subject studied alongside degree performance in the multinomial logit regressions should mitigate the size of bias. In fact, each of these factors is expected to capture some of the differences in individual effort levels.

more likely to be inactive.[24] Mature students are more likely be unemployed or to study for a postgraduate degree and less likely to take up a temporary occupation than their younger peers. This result may reflect, on one hand, the negative signal that a late degree gives to employers. On the other hand, the job search for mature students may be lengthier as they are expected to be more selective in filling applications and evaluating job offers. Not surprisingly, part-time graduates (PTIME) are 20 percentage points more likely to find permanent employment than their full-time counterpart. Part-timers are usually working students, who can capitalise on their work experience, and also on a better knowledge of the labour market. Part-time graduates may also have stronger incentives to find work;[25]

vi. *Other schooling and course characteristics.* Graduates with A-level Mathematics (ALMATH) are nearly 4 percentage points more likely to find permanent employment, even after controlling for degree subject studied. This result confirms to an extent previous evidence of a positive impact of A-level Mathematics on earnings (Dolton and Vignoles, 1999; Naylor *et al.*, 2002). Interestingly, course duration (CDUR) and vocational entry qualifications (OENTQ1) are both found to have a significant and positive impact on the probability of permanent employment, even after controlling for subject studied, type of institution and part-time status;

vii. *Department-related factors.* Table 2.3 also shows interesting departmental effects. Graduates from departments with higher proportions of first and upper second class honours (DP211) are significantly less likely to enter permanent

24 The reliability of these effects is however questionable, due to the possible endogeneity of marriage decisions.

25 Besides, some part-time graduates need not even seek work, as they are likely to be promoted to a higher professional level by their current employer.

employment and more likely to enter temporary employment or unemployment. More precisely, the individual's probability of finding a permanent job is reduced by 1 percentage point for any 10 percentage points increase in the proportion of 'good' graduates within his or her department. This is not counterintuitive if one considers that, on average, a higher number of good graduates from the same department, may have the effect of raising the level of the competition for permanent employment posts. Not surprisingly, a higher proportion of postgraduate students (DPPG) as well as a higher staff/student ratio (DPSSR) in the department increase the probability of postgraduate study. A high staff/student ratio can imply better supervision prospects, while a high number of postgraduates could signal the high quality of the postgraduate courses offered by the department. Furthermore, undergraduate finalists in departments running large postgraduate programs are more exposed to different aspects of postgraduate education, which may exert a positive influence on their choices to stay on. Finally, in line with the results discussed earlier in this section, an increase of 10 percentage points in the proportion of male graduates reduces the probability of employment by 1 percentage point. These significant gender effects motivate the analysis that follows.

2.7 A gender analysis

The results for all students commented in the previous section pointed to significant gender differences in first destination behaviour. Female graduates are more likely than males to engage in vocational training and less likely to enrol in postgraduate courses. Females are also less likely to be unemployed and more likely to be inactive. In this section we re-estimate our multinomial logistic model separately for females (29,391 individuals) and males (34,315 individuals). This will have the effect of removing the constraint that individual characteristics affect first destination behaviour of men and women in the same way. The effect of marital status discussed earlier is a clear example of how significant gender differences can emerge if those restrictions are relaxed.

Summary statistics presented in Table 2.4 show that men and women tend to differ in their schooling curricula, academic performance and in their preferences for particular degree courses. For instance, men dominate in Engineering, Computer Science, Physical Sciences, and Economics. On the other hand, women clearly dominate in Life Sciences, Allied Medicine, Sociology, Classics, Modern European Languages, and Education. Dolton and Makepeace (1990) also found similar results for a sample of 1980 graduates in the UK. In terms of degree performance, more males obtain first class honours but female graduates perform better if one considers the aggregate proportion of 'good' degrees (first and upper second class). Similarly, more males obtain 28 points or higher in the sum of the best three A-level passes (SCOREA). More males take A-level Mathematics (44.5% versus 26.3% for females), hold vocational entry qualifications (OENTQ1) and graduate from EXCAT institutions.

Table 2.5 reports the marginal effects derived from the gender-specific equations. Results are only shown for those characteristics where significant gender differences emerge, while the full set of results can be found in Appendix 2B. The main findings are summarised below:

i. The 'gradient' of degree performance is much steeper for males, especially in terms of the probability of unemployment. Although the magnitude of the effects may be partly driven by differences in sample variability (the distribution of degree class is more disperse in the male sample), the results seem to suggest that the penalty of getting a 'poor' degree (lower second class or lower) is significantly larger for males;

ii. Relative to those from intermediate parental backgrounds, graduates coming from partly skilled or unskilled families are significantly less likely to find permanent employment, but this is only true for males. Likewise, SC IV-V graduates are more likely to be unemployed, especially if male. This may indicate that social networks defined as social ties to those in high-paying jobs are important, particularly in typical male occupations;

iii. Interestingly, having attended an independent school (INDEP) significantly improves the probability of permanent employment, particularly for males, after controlling for parental social class. Private schools may instil certain social skills valuable to employers and/or they may benefit pupils indirectly through affiliation to particular business and social networks unaccounted for by the occupational status of parents;

iv. In line with the results in the previous section, marital status (MARRIED) has a very different impact on early career patterns of male and female graduates. Married men are 9 percentage points more likely to enter a permanent occupation and 5 percentage points less likely to be unemployed than their single counterparts. As for females, being married reduces the probability of permanent employment by 3.5 percentage points and positively affects the probability of being inactive (+4 percentage points);

v. A higher departmental proportion of postgraduates (DPPG) significantly increases the probability of further study, but only for males. It is also interesting to notice that a higher proportion of male students (DPMALE) significantly increases the probability of postgraduate study for male and female graduates;

vi. Finally, we find interesting gender differences in the effects of subject studied. Whilst, relative to Economics, reading Engineering or Computer Science significantly increases the probability of entering a permanent occupation for males the corresponding effects are absent among females. Likewise, reading Classics, Modern European Languages, and Humanities significantly increases the probability of unemployment only for men. For instance, a female with a degree in Modern European Languages is 2.2 percentage points less likely to be unemployed than the observationally equivalent female with a degree in Economics. On the other hand, a male with a Language degree is 4.1 percentage points more likely to be unemployed. These results seem to indicate that female (male) graduates tend to perform worse than males (females) in traditionally male-dominated (female-dominated) subjects, after controlling for prior qualifications and degree performance. This, in turn, may explain the different concentrations of female and male students across the spectrum of academic subjects. Students appear to choose those courses that are perceived to maximise their employability upon graduation, conditional on their endowments of other productivity-related characteristics at the time in which the decision is made.[26] Similar evidence on comparative advantage has been used in the related literature on the gender wage gap to explain field choice and differences in field concentration by gender (Paglin and Rufolo, 1990; Brown and Corcoran, 1997).

2.7.1 Further analysis on gender differences

Individual perceptions can be based on students' awareness of their own strengths and weaknesses (for instance, someone with undistinguished numerical skills will decide not to choose scientific courses), but may also reflect awareness of external factors like entry barriers to some jobs (Anker, 1997; Preston, 1999) or glass-ceiling effects (Dolton *et al.*, 1996). For instance, women may decide to stay out of traditional 'male' courses because they perceive that the labour market would reward them less for these courses than men. To gain an insight on these issues, we decompose gender differences in first destination behaviour into two components: the portion explained by differences in characteristics that we can directly observe and control for, and a residual portion which is ascribed to unobservable factors, including possible labour market discrimination. To this end, we draw upon a methodology commonly used in the discrimination literature to analyse the gender wage gap (Oaxaca, 1973; Blinder, 1973; Daymont and Andrisani, 1984).

More formally,

---

26 A simultaneous model of first destination decisions and subject choice would be preferred to properly account for the endogeneity of the degree course. This is beyond the scope of this chapter. However, we will return to this issue in Chapter 5 when we estimate a simultaneous model of earnings determination and subject choice.

$$P_j^m - P_j^f = \left[\frac{1}{M}\sum_{i=1}^{M}\hat{p}_{ij}^m - \frac{1}{F}\sum_{i=1}^{F}\hat{p}_{ij}^m\right] + \left[\frac{1}{F}\sum_{i=1}^{F}\hat{p}_{ij}^m - \frac{1}{F}\sum_{i=1}^{F}\hat{p}_{ij}^f\right] \tag{2.5}$$

or, alternatively, if individual female probabilities are used as standard,[27]

$$P_j^m - P_j^f = \left[\frac{1}{M}\sum_{i=1}^{M}\hat{p}_{ij}^f - \frac{1}{F}\sum_{i=1}^{F}\hat{p}_{ij}^f\right] + \left[\frac{1}{M}\sum_{i=1}^{M}\hat{p}_{ij}^m - \frac{1}{M}\sum_{i=1}^{M}\hat{p}_{ij}^f\right] \tag{2.6}$$

where $\hat{p}_{ij}^f$ ($\hat{p}_{ij}^m$) are the individual predicted probabilities that the female (male) graduate i chooses alternative j, P_j^f (P_j^m) are the observed female (male) probabilities that outcome j is chosen ($j=1$), and F (M) is the total number of female (male) graduates in the sample. The first component of (2.5) and (2.6) can be interpreted as the percentage of the gender gap in the probability that outcome j is chosen due to differences in sample mean characteristics, whilst the second component captures the residual part of the gap that is left unexplained. The results of these decompositions are shown in Table 2.6. The first column reports the gender gap in the observed first destination probabilities. The larger differences are found with respect to professional training (-8.87 percentage points), postgraduate study (5.67 percentage points) and unemployment (5.52 percentage points). The decomposition of these differences reveals that, while the model explains a substantial proportion of the gender gap in the probabilities of further study and professional training, the unemployment gap remains largely unexplained. This can be due to several concurring factors. It is possible that

27 This alternative decomposition is done for a sensitivity check (Machin and Puhani, 2003).

women are more dedicated and effective in their job search. It is also possible that the labour market is segmented along gender lines and unobserved sex-based differences in the type of occupation and sector of employment could account for a significant part of the gap. A third possibility could be the presence of affirmative-action recruiting and hiring policies favouring females.[28] The decomposition (not shown) of the observed gap in the probability of PWORK seems to validate the conclusion that differences in unmeasured characteristics tend to favour females. Male graduates would be 1.5 percentage points more likely than females to find a permanent job if only differences in observable factors are taken into account. However, men are overall 0.74 percentage points less likely to enter a permanent occupation because of an adverse effect, relative to females, originating from unmeasured factors.

2.8 A subject-specific analysis

A second extension to the analysis presented in Section 2.6 involves exploring whether the effects of some key determinants of graduates' early careers vary depending on the subject studied at university. A subject-specific analysis not only removes the constraint, intrinsic in the previous models, that the effects of variables like degree class, social class and schooling background are equal across different degree courses, but is also an interesting tool to address two important policy issues:

---

[28] The employers' willingness to hire stigmatized applicants may have increased since the early 1980s. The evidence of a growing number of job openings highlighting the equal opportunity nature of the employer and encouraging applications from females candidates seem to confirm this tendency.

i. The debate surrounding the possibility that universities could in the future charge higher tuition fees for some degree courses in high demand;

ii. The reform of the secondary schooling curriculum, and more specifically, the controversial proposal of broadening curricula at A-level. If individuals taking a narrower curriculum at 16-19 also tend to choose certain degree courses, conditioning on degree subject will be important to ascertain the effect of curriculum breadth on graduates' first destinations.

These issues are discussed in turn in Sections 2.8.1 and 2.8.2. The analysis is carried out breaking down the sample by subject studied and estimating Equation (2.2) for each of the following courses: Biology (BIOL), Physical Sciences (PHYS+CHEM), Mathematics (MATHS), Computer Science (COMP), Engineering (ENGIN), Business Studies (BUS), Economics (ECON), Law, Other Social Studies (SOCIO+POL+OSOSCI), Humanities (HUM), Modern European Languages (MEUL), and Education (EDU).[29] Splitting the sample by subject studied necessarily implies a considerable reduction in the number of observations available for each estimation. Cell size considerations have required a more parsimonious model specification. The number of first destination outcomes was restricted to three: i) WORK (PWORK+TWORK), ii) STUDY (PSTUDY+TRAIN), iii) OLFU (OLF+UN).[30] Social class I, II, and IIINM were merged into SCHIGH, while SC IIIM and SC IV-V were aggregated and renamed

---

[29] These aggregations were dictated by cell size considerations and involve courses not too dissimilar in their first destination effects. Besides, aggregation should not undermine interpretation given that the merged-in courses belong to the same subject area.

[30] This restricted specification has also the advantage of making our results more comparable to those from existing studies, which have used a similar three-way classification (McKnight, 1999; Johnes and Taylor, 1990).

into SCLOW.[31] With respect to degree class, we restrict our attention to three groups of graduates: those who achieved a first honours degree class (DFIRST), those who qualified with upper second honours (DUSEC), and those who graduated with lower second class honours or lower (POORDEG). There is widespread perception among both employers and students that a '2.1' is an important threshold that serves as a screening device for academic ability (Smith and Naylor, 2001; Purcell and Pitcher, 1996; Roizen and Jepson, 1985). Finally, female and male graduates in each subject are now considered together to avoid small cell sizes in courses which are dominated by one gender.

2.8.1 Student finance

The current funding system based on a mix of repayable loans and means-tested fees has come under increasing criticism not only by students and their parents but also by academics and Vice-Chancellors. In fact, not only has the system proved inadequate to provide the cash that many under-funded universities need, it has also been flagged as being regressive in the sense of putting off less affluent students.[32] However, while everyone agrees on the inadequacy of the current system of student finance, views differ radically on how to improve it. One solution recently proposed by some academics and also favoured by the

---

[31] Aggregation is justified by earlier evidence of small or insignificant differences between the effects of the merged social groups. SC OTH is entered as a separate category.

[32] The 2003 White Paper '*The Future of Higher Education*' published on 22 January 2003 by the DfES affirms that universities will be able to set their own fees between £0 and £3,000 by 2006, provided they satisfy the access regulator appointed *ad hoc* to ensure that students from disadvantaged backgrounds are adequately represented. Students will get a grant of £1,000 a year if their parents earn less than £10,000, and there will be no increase on interest on loan repayments, which will increase only by the level of inflation. Furthermore, payback will be deferred and will start when the graduate earns more than £15,000.

University UK (UUK),[33] the body representing Vice-Chancellors, is to allow universities to charge more for popular courses.[34] However, the introduction of 'market fees' is controversial. Critics fear that top-up fees could negatively affect equal opportunities and intergenerational mobility, insofar as students from less affluent backgrounds may be diverted away from courses which are more expensive but that, in return, offer better career prospects.[35]

One way to address the issue is to suppose that when choosing which university course to enrol in students operate a trade-off between the return and risk associated with their choice (Mingat and Eicher, 1982). First, let us define as 'high-return' a course associated with a relatively high probability of employment upon graduation.[36] Then, let us define as 'high-risk' a course where the penalty that the labour market inflicts for a poor academic performance is comparatively high because the 'cream-skimming' of the talent is more pervasive. If market or top-up fees are introduced, it is plausible to expect that high-return courses will become relatively more expensive, because of demand-pull effects on prices.[37] If high-return courses are, on average, high-risk, risk-averse students may decide to take subjects that are potentially less rewarding but represent a safer investment

---

33 The Committee of Vice-Chancellors and Principals (CVCP) was renamed Universities UK (UUK) on 1 December, 2000.

34 Staff and Agencies: 'Universities could charge more for popular courses'. The Guardian, 26 September, 2001.

35 Tony Higgins: 'An unfair reflection'. The Guardian, 11 January, 2002

36 Usually the word 'return' is used with reference to the financial benefits, such as occupational earnings or starting salaries. As we will show in Chapter 5, there is a clear link between subjects associated with high employability and subjects that command high financial returns.

37 Even if higher education will not evolve into a full-fledged market in the sense that university fees will not be determined solely by demand and supply factors, it is likely that high-return courses will nonetheless be relatively more expensive. In fact, equity considerations could require students taking high-return courses to contribute more, as the main beneficiaries of these returns,

option. If one accepts that risk-aversion is linked to socio-economic factors, differential fees could distort students' choices in a way that adversely affects intergenerational mobility, at least in the absence of fee exemption mechanisms or subsidies to individuals from less privileged backgrounds.

To test empirically the validity of these predictions, we construct return and risk indicators for each subject. After estimating three-way (WORK, STUDY and OLFU) subject-specific multinomial logit models of first destination, the predicted probability of WORK (RET1) or, alternatively, 1 minus the predicted probability of OLFU (RET2) are used as indicators of return. To measure risk we calculate, for each subject, the marginal effect of graduating with a '2.2' or lower (POORDEG) relative to a '2.1' on the probability of WORK (RISK1) or, alternatively, on the probability of OLFU (RISK2).[38]

Summary statistics (Table 2.7) show that the distribution of students by degree class varies considerably according to subject studied. A relatively high proportion of 'low achievers' (POORDEG) is observed in Education, Mathematics, Physics, Computer Science, and Engineering. Interestingly, Mathematics, Physics, Computer Science, and Engineering are also the courses with the highest concentration of students from less privileged backgrounds (SCLOW).

---

towards overall tuition costs. Such policy would reflect the same principle that inspired the introduction of flat-rate university tuition fees in the autumn of 1998 in the UK.

[38] According to the RISK1 (RISK2) indicator, the riskier subjects are those with high negative (positive) marginal effects.

The results from the multinomial logit equations estimated for each of the subjects defined in Section 2.8 are reported in Table 2.8. With respect to RET1, the highest scoring subject is Education, followed by Business Studies, Computer Science, Engineering, and Economics. When subjects are ranked according to RET2, Law graduates rank first, followed by Education, Modern European Language Business and Physical Sciences.

With respect to risk, the impact of degree performance on first destination outcomes is always statistically significant and varies markedly between subjects. In particular, POORDEG has a negative impact on the employment prospects of engineers, mathematicians, computer scientists, economists and Business graduates, after controlling for institution attended, gender and prior qualifications. In these courses, POORDEG significantly depresses the probability of WORK (RISK1) on one hand, and enhances the probability of OLFU (RISK2) relative to the 'otherwise observationally identical' individual with a '2.1', on the other hand. Interestingly, POORDEG has positive effects on the predicted probability of WORK for graduates in most of the Science (Biology, Physics) and Art (Humanities, Modern European Languages and Other Social Studies).disciplines. We also note that RISK2 is significantly lower in Art subjects and almost zero in Education degrees.

Table 2.8 also suggests that Law graduates from a working class background (SCLOW) are significantly more likely to be unemployed than well off graduates. The influence of social and business networks may be particularly strong in legal professions. The fact that over 16% of all graduates who have parents in legal occupations read Law, lend support to the idea that parental and social background have a significant impact on the choice to pursue certain careers.[39] This, in turn, may explain why relatively few working class students take Law. We also find some interesting independent schools' effects (INDEP). Graduates in Biology who went to private schools are more likely to find employment relative to 'otherwise observationally identical' individuals who went to LEA schools. On the contrary, a negative and significant effect is found for Law graduates.

To ascertain whether and to what extent 'high-return' courses also tend to be 'high-risk' we ranked subjects according to the indicators of risk and return discussed above and then calculated the correlation coefficients between pairs of rankings. Table 2.9 shows a positive correlation coefficient of 0.84 between rankings based on RET1 and RISK1. This means that those subjects where the probability of entering an occupation upon graduation is relatively high, also tend to be the courses where the chances of finding a job are more sensitive to degree performance. The estimated coefficient is also statistically significant at the 1% level, unlike the results obtained when alternative combinations of the risk-return indicators defined above are considered. Before discussing the potential implications of this correlation analysis, one should be cautious about the fact that for some subjects the negative signal to employers of achieving a low degree class may be 'louder' partly because the pool of low achievers is smaller. Put differently, the indicator RISK1 is 'conditional' on having obtained a '2.2' or lower. Therefore, the reliability of RISK1 would be called into question if subjects where employability is highly sensitive to degree performance are also those

39 Chevalier (2002) suggests the presence of nepotism for some subjects in the UK, including Law and Medicine.

where graduates are, on average, less likely to graduate with a low class degree.[40] To address this potential weakness of the RISK1 indicator, we calculate an 'unconditional' indicator of risk (RISK3) as the product between RISK1 and the observed proportion of low achievers shown in Table 2.7.[41] Table 2.9 shows a positive correlation coefficient of 0.94 between RISK1 and RISK3, and a positive correlation coefficient of 0.75 between RET and RISK3. Both values are statistically significant at the 1% level. These results indicate that high-return subjects also tend to be high-risk.

The risk-return analysis presented in this section has potential implications for the ongoing debate on student finance. The popularity of Engineering, Computer Science, Business and Economics degrees as courses which tend to offer good employment and pay prospects is likely to make these subjects prime candidates for 'top-up' fees (Naylor *et al.*, 2002). However, as they become more expensive, risk-averse applicants may decide to choose 'safer', but ultimately less remunerative degrees. This need not be necessarily true if an adequate cross-subsidisation mechanism is devised to balance out the higher opportunity costs faced by less affluent students. However, we believe that further research is required to explore the social impact of alternative funding options before market fees are introduced.[42]

---

[40] This may be due to the fact that some subjects are academically less demanding or to differences in awarding standards.

[41] We also estimated subject-specific probit models of degree performance to predict the probability of achieving a 'poor' degree. The regressions controlled for a wide range of personal and academic characteristics including A-level subjects, count and grade score, type of school attended, social class, gender, age, residence, course duration and some departmental variables. We found that the predicted probabilities of POORDEG were very similar to the 'unadjusted' POORDEG proportions commented in Table 2.7, so we decided not to report the results.

[42] The results may suffer from a selection bias originating from the endogeneity of subject studied, in that individuals 'self-select' into courses where they have a higher-than-average probability to succeed. We note, however, that the bias is likely to mitigate both the differences in degree

2.8.2 Secondary school curriculum

Critics claim that the English secondary school system, and in particular the A-level curriculum, inadequately prepares students for the world of work. One recurrent argument is that the curriculum for 16-19 year olds is too narrow and provides overly specialised knowledge. In response to this criticism, the government has recently committed itself to broadening the curriculum at 16-19. Since September 2000, pupils are expected to take up to 5 different subjects in their first year of sixth form as well as key skills courses in information technology, communication and 'application of number', while only in their last year of secondary schooling will they specialise taking the normal three A-level subjects.[43]

With the exception of A-level Mathematics, previous empirical studies on the effects of different types of schooling curricula on graduate earnings have generally found little evidence that employers reward specific subjects (Dolton and Vignoles, 1999; Naylor *et al.*, 2002; Altonji, 1995). Dolton and Vignoles (2001) investigate the effect of curriculum breadth on earnings. Using information on individuals who graduated in 1980 in the UK, they conclude that employers do not seem to reward individuals who take a broader curriculum more highly. This conclusion proved robust to different measures of curriculum breadth including the number of different A-level specialisations taken (Science-related, Social Sciences-related, or Art-related), whether an individual had taken the General Studies A-level, or whether the individual is Scottish (this is a 'natural

performance and the differences in employability across disciplines. This means that the estimated differences in graduate employability across subjects are expected to be a lower bound measure of the true differences that would be observed if self-selection was taken into account.

43 DfEE Press Notice, 3 April 1998.

experiment' given that in Scotland 5 SCE Highers are the norm compared to 3 GCE A-levels for British students). Furthermore, results did not change even when first job's annual earnings were used instead of 1986 earnings (time of the survey). In fact, one would expect employers to extract a stronger signal from the secondary school curriculum of newly qualified rather than more experienced graduates.

However, other studies based on surveys of graduates and employers have highlighted how employers are beginning to see the skills instilled by Art disciplines such as the ability to communicate and interact with others as essential and integral elements in the competence of effective professional practitioners in scientific, technical and other fields (Squires, 1990; Lynton, 1993).

In view of this seemingly conflicting evidence, and in consideration of the policy relevance of these issues, this section investigates whether A-level curriculum, and in particular curriculum's breadth, has any impact on the employability of newly qualified graduates.[44] As in Dolton and Vignoles (2001), we use two measures of curriculum breadth: the number of British A-levels or Scottish Highers that each student took in different areas of specialisation, and whether or not the student took the General Studies A-level. To construct the first indicator four alternative groups of disciplines are considered: Science, Social Sciences, Humanities, and Other (including other minor subjects) (see Appendix 2A for details). The variable CURR takes values between 0 and 3, depending on whether the student took none (non-A-level entry qualifications), one, two, or three or more different specialisations. CURR is entered in the analysis as a categorical

[44] It can also be the case that a narrow curriculum at A-levels offers a competitive advantage in some technical occupations.

variable to account for the potential non-linear relationship between curriculum breadth and first destination outcomes. The rationale for the inclusion of the second indicator of curriculum breadth, namely whether or not the student took the A-level General Studies, stems from the fact that this subject was introduced in 1959 with the intent of broadening pupils' curriculum at 16-19, and it has now become the second most studied subject (Dolton and Vignoles, 2001). A dummy variable ALGENS is included in the regressions to account for this effect.

To assess the effect of curriculum breadth on graduates' first destinations we estimate Equation (2.2) for each of the 13 subjects defined in Section 2.8. To the usual set of controls and the two measures for curriculum breadth discussed above (ALGENS and CURR), a dummy variable (JOINTDEG) for graduates who took joint degrees was also added to the regressions in order to disentangle the effect of curriculum breadth at 16-19 from curriculum breadth at university. It is worth noting that controls for A-level count (ALCOUNT and HCOUNT) and average score are particularly important because curriculum breadth can be endogeneous if more able students also have a broader curriculum.[45] Finally, as in the previous regressions, we include a separate dummy for A-level Mathematics in view of its significant effects on graduate earnings found in related studies.

[45] We note that General Studies A-levels are traditionally taken in addition to the three base A-level subjects required to gain entry in many institutions. Usually, more able students take this subject. Therefore, ALGENS can also be considered as a proxy of ability. This implies that it may be difficult to extract the 'curriculum breadth' effect of this variable. On the other hand, the inclusion of ALGENS alongside CURR in the regression is likely to improve the quality of the latter as a proxy for curriculum breadth. I thank Prof. Shackleton for this suggestion.

Descriptive statistics (Table 2.7) show that graduates in Science disciplines and in Engineering have typically a narrower A-level curriculum. The wider curricula (CURR3) are found among economists, other social scientists and Business graduates. Interdisciplinary degrees (JOINTDEG) are more popular among social scientists (excluding Law) and mathematicians and less popular among engineers, biologists and Physics graduates. Finally, the proportion of graduates with General Studies A-levels (ALGENS) is around 20% across the board. In terms of the overall number of A-levels, graduates in Mathematics and Engineering take on average more A-levels than individuals from other courses, while for Scottish students, relatively high Higher counts are found among graduates from Economics, Business and Law.

Table 2.10 shows the marginal effects of the curriculum-related variables on first destination by degree course. A broader curriculum significantly and monotonically increases the probability of employment for graduates in Biological Sciences, Computer Sciences and Engineering after controlling for A-level count and grade scores. No effects are found in other courses. With respect to ALGENS, graduates in Biology, Physical Sciences, Economics and Business who took this subject at 16-19 have a higher probability of employment, while for Language graduates A-level General Studies reduce the probability of unemployment.

Taken together, these results seem to suggest that broader curricula enhance employability upon graduation, particularly in some technical degree courses like Engineering and Computer Science where curricula tend to be relatively narrow.

These results seem to support the view that employers may value a technical degree more highly if graduates have some breadth of knowledge cutting across different areas of specialisation. Our findings possibly complement, rather than contradict, Dolton and Vignoles' conclusion of no financial returns to a broader A-level curriculum for newly qualified graduates. The fact that a broader curriculum does not seem to command higher starting salaries need not deter students from taking a larger variety of subjects at A-levels. In fact, starting salaries are often a poor indication of future career earnings. On the other hand, students may perceive the extra effort made to broaden their curriculum as worthwhile if this facilitates their transition from university to work. Finally, the evidence produced by Dolton and Vignoles refers to a sample of 1980 graduates. The same may not be true for later cohorts. The positive signal that a broad curriculum conveys to employers may have grown stronger in recent years as a result of a tighter and more fragmented labour market for the highly qualified. Individuals with some degree of interdisciplinary knowledge can be more easily trainable to perform different tasks and this can give them an edge in an increasingly competitive working environment.

Finally, A-level Mathematics (ALMATH) improves the probability of entering an occupation for Economics, Law, Humanities, Languages and Education graduates. With the exception of Economics, these subjects have typically low proportions of students with this specialisation. Therefore, from a policy perspective, it seems important to encourage more students to take this A-level, regardless of their field of study at university. This result is in line with previous evidence suggesting that graduates with A-level Mathematics tend to receive higher earnings.

2.9 Summary and conclusions

The chapter has looked at the first destination choices of newly qualified graduates using individual-level USR data on the complete cohort of 1993 university leavers in the UK. We considered an unusually broad range of alternative destinations including permanent work, temporary work, vocational training, postgraduate study, unemployment and non-employment and estimated the probabilities that each of the above destinations is chosen, conditional on a wealth of personal and university-related characteristics including gender, social class, subject studied, university attended, degree performance, and secondary schooling curriculum. We found a number of interesting results, some of which have potentially important policy implications:

i. Female graduates are more likely than males to engage in vocational training or to be inactive and less likely to enrol in postgraduate courses or to be unemployed. We found that these differences, and especially the unemployment gap between sexes, are largely due to gender differences in unobservable factors. A speculative explanation behind the lower probability of unemployment among females could be related to unobserved changes in recruiting and hiring policies aimed at encouraging job applications from female graduates;

ii. Students from working class backgrounds (skilled manual, partly skilled and unskilled parents) are more likely to be unemployed or inactive six months after graduation, especially if male. This may indicate that social networks defined as social ties to those in high-paying jobs are important, particularly in typical male occupations. The potential role of social networks in facilitating the transition from university to the world of work is confirmed by additional evidence from the subject-specific analysis. We find that working class graduates in Law are significantly more likely to be unemployed than their wealthier peers. The influence of social and business networks may be particularly strong in legal professions, especially if one considers the relatively high proportion (over 16%) of Law graduates whose parents are in legal occupations (judges, magistrates, solicitors, barristers, and advocates). 'Inside tracks' working through a system of job referrals may provide an explanation to the observed effects;

iii. In some subjects the penalty that the labour market inflicts for a poor degree performance is relatively high. If this penalty can be regarded as a measure of risk, some subjects are relatively riskier that others. We also find that riskier (so defined) subjects offer better career prospects upon graduation. These results have interesting policy implications, which are linked to the heated debate on the reform of student finance in the UK. Because, despite the hazard, risky subjects are popular among students, their relative price is likely to increase if fees are liberalised. A rise in tuition costs will exacerbate risk, because degree performance is unpredictable *a priori*. Assuming that working class students are more risk-averse, fee liberalisation can create a particularly strong deterrent for these students from choosing riskier but ultimately more economically rewarding degrees, with negative consequences on intergenerational mobility. Although an assessment of alternative funding options is beyond the scope of this Thesis, we believe that further research is required to explore the social impact of market fees under alternative regimes of student finance;

iv. A broader secondary school curriculum enhances employability upon graduation, especially for graduates in Engineering and Computer Science. These results seem to support the view that employers tend to value technical degrees more highly if graduates have some breadth of knowledge cutting across different areas of specialisation. These findings are informative of the current debate on reforming A-level curricula in the UK. In particular, evidence suggesting that a broader A-level curriculum does not command higher starting salaries (Dolton and Vignoles, 2001) need not deter students from taking a larger variety of subjects at A-levels. In fact, students may perceive the extra effort for a more rounded and differentiated curriculum worthwhile if this facilitates their transition from university to work;

v. A-level Mathematics improves the probability of entering an occupation for Economics, Law, Humanities, Languages and Education graduates. This result supports previous evidence suggesting that A-level Mathematics positively affects future earnings (Dolton and Vignoles, 2001; Naylor *et al.*, 2002). With the exception of Economics, these subjects have typically low proportions of students taking Mathematics at A-level. Therefore, from a policy point of view, it seems important to encourage more students to take this subject, regardless of their field of study at university.

The next chapter extends the first destination analysis over time by considering fourteen different cohorts of graduates. This enables us to gauge possible changes in first destination behaviour over time. The rapid growth in participation rates in higher education on one hand, and a tightening and increasingly fragmented marketplace for the highly qualified on the other, have enhanced the screening function that employers attach not just to any degree but rather to the most selective and demanding programs as well as to the most prestigious institutions.

Therefore, it would be interesting to examine how the effects of the university attended and the subject studied on first destinations have evolved over time. Moreover, the distinction between permanent and temporary work is probably too simple to address the important issue of how personal and educational characteristics influence the probability of graduates to enter not just any occupation, but one where the skills and knowledge learnt at university are a prerequisite for the job. In the next chapter we will extend our first destination analysis by using a finer distinction between 'graduate' occupations (requiring a first degree) and 'non-graduate' occupations for which graduates are typically overqualified.

Tables and Figures

Table 2.1 Likelihood ratio tests for category pooling [a]

categories [b]	χ^2	degrees of freedom	$p>\chi^2$
2-1	1137.8	62	0.00
2-3	1942.8	62	0.00
2-4	2916.1	62	0.00
2-5	409.2	62	0.00
2-6	425.4	62	0.00
3-1	7197.1	62	0.00
3-4	6920.7	62	0.00
3-5	3871.8	62	0.00
3-6	1371.5	62	0.00
4-1	6002.2	62	0.00
4-5	4208.1	62	0.00
4-6	1969.6	62	0.00
5-1	2203.3	62	0.00
5-6	884.5	62	0.00
6-1	899.1	62	0.00

[a] H_0: All coefficients except intercept associated with a given pair of outcomes are equal to zero.

[b] 1=PWORK; 2=TWORK; 3=TRAIN; 4=PSTUDY; 5=UN; 6=OLF.

Table 2.2 Descriptive statistics for all students (%)

variable	respond	non-respond	variable	respond	non-respond
PWORK	44.7	-			
TWORK	5.9	-			
TRAIN	16.0	-			
PSTUDY	16.9	-			
UN	11.5	-			
OLF	5.0	-			
BIOL	5.1	3.7	SC I	18.3	15.1
OBIOL	4.4	4.7	SC II	42.3	41.6
CHEM	3.8	2.6	SC IIINM	11.2	10.2
PHYS	3.3	2.4	SC IIIM	10.1	7.8
OPHYS	3.7	2.9	SC IV-V	6.8	6.9
MATHS	5.2	3.4	DFIRST	10.3	4.4
COMP	3.2	3.1	DUSEC	49.0	40.0
ECON	5.2	6.3	DLSEC	31.3	38.5
SOCIO	2.9	4.9	OCIV	34.8	34.9
POL	2.9	4.1	OXBR	7.3	7.4
LAW	5.2	3.8	NCIV	17.4	19.6
OSOSCI	3.7	4.2	EXCAT	13.9	15.0
CLAS	7.3	9.2	NEW60	16.1	15.0
ARTS	1.9	2.8	INDEP	23.4	24.5
MEUL	4.9	5.2	MATURE	11.3	20.4
ALMED	3.6	1.9	FEMALE	46.1	41.3
ENGIN	10.5	9.5	MARRIED	2.9	3.7
BUSIN	5.9	4.3	ALMATH	45.8	34.6
HUM	8.1	10.8	PTIME	1.6	2.0
EDU	1.7	1.1	OENTQ1	3.5	3.0

Table 2.3 Multinomial logit estimates: marginal effects (all students)

variable	PWORK	TWORK	TRAIN	PSTUDY	UN	OLF
BIOL	-0.212*	-0.002*	-0.001	0.206*	0.002	0.006*
CHEM	-0.272*	-0.027	0.028*	0.317*	-0.030*	-0.017*
PHYS	-0.254*	-0.012	0.082*	0.195*	0.009	-0.021
COMP	0.098*	-0.020	-0.063*	-0.014	0.020	-0.020*
LAW	-0.396*	-0.039*	0.586*	-0.059*	-0.079*	-0.013*
CLAS	-0.184*	0.007	0.123*	0.020	0.029	0.004
ALMED	0.130*	-0.033*	-0.090*	0.069*	-0.049*	-0.026*
MEUL	-0.157*	-0.006*	0.199*	-0.030*	0.004*	-0.010
ENGIN	0.037*	-0.010*	-0.027*	0.018*	-0.002	-0.016*
BUS	0.079*	0.005	-0.012	-0.066*	-0.013	0.007
SOCIO	-0.114*	0.021*	0.031*	0.015*	0.034*	0.014*
POL	-0.129*	0.010*	0.081*	0.031*	0.013*	-0.005*
HUM	-0.164*	0.003	0.120*	0.024*	0.019*	-0.003
EDU	0.152*	-0.011*	-0.021*	-0.087*	-0.038*	0.005
OXBR	0.065*	-0.021*	0.034	-0.018	-0.039*	-0.020*
EXCAT	0.074*	0.008	-0.022	-0.037*	-0.006	-0.016*
DFIRST	-0.019*	-0.030*	-0.035*	0.149*	-0.055*	-0.010
DLSEC	0.003*	0.026*	-0.001*	-0.097*	0.059*	0.011
DTHIRD	-0.044*	0.037*	0.013*	-0.130*	0.113*	0.011
SC I	-0.025	-0.001	0.009	0.020*	-0.006	0.004
SC IV-V	-0.028*	0.017*	-0.012	-0.001	0.033*	-0.010
SCOREA	0.081*	-0.014*	-0.019*	-0.041*	-0.016*	0.009
SCOREB	0.065*	-0.007	-0.009*	-0.041*	-0.009	0.003
SCOREC	0.029*	-0.001	0.002	-0.030*	0.000	0.001
ALMATH	0.037*	-0.005	-0.023*	0.003	-0.009	-0.003
OENTQ1	0.073*	0.004	-0.018	-0.035*	-0.013*	-0.011
CDUR	0.101*	-0.008*	-0.048*	-0.020*	-0.019*	-0.007*
PTIME	0.194*	-0.038*	-0.053*	-0.026*	-0.054*	-0.026*
FEMALE	0.031*	0.002*	0.049*	-0.035*	-0.053*	0.006*
MARRIED	0.083*	-0.017*	0.019*	-0.020*	-0.043*	-0.021
FEM*MARR	-0.138*	-0.009*	-0.031*	0.009	0.055*	0.113*
DPMALE	-0.081*	0.008	-0.060*	0.096*	0.019*	0.017*
DPR211	-0.100*	0.037*	0.007*	-0.049*	0.050*	0.055*
DPSSR	-0.280*	-0.094*	0.044*	0.378*	-0.087*	0.039*
DPPG	-0.072*	-0.009*	0.032*	0.047*	0.011*	-0.009*
N	63706					
LL	-86352					
pseudo R^2	0.11					

* denotes statistical significance at 5% level.

Table 2.4 Descriptive statistics by gender (%)

Variable	females	males	variable	females	males
PWORK	45.1	44.3	SOC I	18.3	18.2
TWORK	6.2	5.7	SOC II	43.1	41.5
TRAIN	20.8	11.9	SOC IIINM	10.9	11.5
PSTUDY	13.8	19.5	SOC IIIM	9.4	10.7
UN	8.5	14.0	SOC IV-V	6.3	7.2
OLF	5.5	4.5	DFIRST	8.3	12.0
BIOL	6.2	4.2	DUSEC	55.0	43.9
OBIOL	6.1	2.9	DLSEC	30.3	32.1
CHEM	2.9	4.5	DTHIRD	6.4	12.0
PHYS	1.3	5.1	OCIV	35.0	34.6
OPHYS	3.1	4.2	OXBR	6.6	7.9
MATHS	4.1	6.0	NCIV	17.9	17.0
COMP	0.8	5.2	EXCAT	11.4	15.9
ECON	3.4	6.8	NEW60	17.4	15.0
SOCIO	4.4	1.7	LEA	47.1	47.8
POL	2.5	3.3	INDEP	22.2	24.4
LAW	5.8	4.6	GRAM	12.4	10.9
OSOSCI	4.6	2.9	SCOREA	18.2	24.0
CLAS	10.8	4.4	SCOREB	24.5	22.6
ARTS	2.7	1.3	SCOREC	25.8	23.0
ALMED	8.0	2.3	SCORED	31.5	30.2
MEUL	5.4	2.0	MARRIED	3.6	2.2
ENGIN	3.0	16.9	ALMATH	35.1	54.9
BUS	5.8	5.9	MATURE	11.0	11.5
HUM	8.5	7.7	PTIME	1.68	1.57
EDU	2.8	0.8	BTEC	2.31	4.49

Table 2.5 Gender-specific predicted probabilities and marginal effects: main differences [(a)]

	females	males
PP (PWORK)	48.4	48.8
PP (TWORK)	6.2	5.5
PP (TRAIN)	19.0	9.5
PP (PSTUDY)	11.6	17.4
PP (UN)	8.7	13.9
PP (OLF)	6.1	4.9
BIOL (TRAIN)	2.3*	-1.3*
CHEM (TRAIN)	10.0*	-0.9*
OPHYS (TRAIN)	9.3*	-0.5*
COMP (PWORK)	1.4	10.5*
SOCIO (UN)	0.1	6.3*
CLAS (UN)	-0.8*	8.0*
MEUL (UN)	-2.2*	4.1*
ENGIN (PWORK)	-0.2*	4.5*
HUM (UN)	0.0*	4.2*
DFIRST (UN)	-2.6*	-7.7*
DUSEC (UN)	3.8*	8.0*
DTHIRD (UN)	6.2*	15.3*
SC IIISM (UN)	0.2	3.0*
SC IV-V (PWORK)	0.1	-3.7*
SC IV-V (UN)	2.1*	3.6*
INDEP (PWORK)	1.1*	4.5*
MATURE (TRAIN)	-4.4*	0.1
MATURE (PSTUDY)	3.7*	0.1
MARRIED (PWORK)	-3.5*	9.2*
MARRIED (UN)	-0.2	-4.7*
MARRIED (OLF)	4.1*	-1.4*
DPPG (PSTUDY)	0.6	8.4*
DPMALE (PSTUDY	10.8*	7.4*
N	29391	34315
LL	-40024	-45955
pseudo R^2	0.09	0.11

[(a)] Predicted probabilities are calculated at sample means.
* denotes statistical significance at 5% level.

Table 2.6 'Oaxaca-Blinder' decompositions of gender differences.

j	$P_j^m - P_j^f$	% gap explained	
		M[a]	F[b]
PWORK	-0.74	43.0	41.0
TWORK	-0.56	24.0	84.0
TRAIN	-8.87	41.0	45.0
PSTUDY	5.67	32.0	47.0
UN	5.52	3.0	11.0
OLF	-1.02	37.0	23.0

[a] Male probabilities used as standard.

[b] Female probabilities used as standard.

Table 2.7 Descriptive statistics by subject studied

	BIOL	PHYS	MATH	COMP	ENGIN	BUS	ECON	LAW	OSOSCI	HUM	CLAS	MEUL	EDU
N	3255	4524	3287	2032	6683	3731	3329	3287	6102	5141	4678	3150	1090
DFIRST	10.8	20.9	18.8	13.7	14.6	7.8	6.9	6.5	5.2	8.6	10.3	6.5	4.6
DUSEC	51.6	32.1	29.7	37.3	35.2	55.9	54.0	60.8	58.0	59.2	58.2	55.8	45.0
POORDEG	37.6	47.1	51.5	49.0	50.2	36.3	39.1	32.7	36.7	32.3	31.5	37.7	50.5
SCHIGH	73.5	70.1	71.7	64.7	67.6	69.2	75.6	77.5	70.8	74.8	76.1	78.6	50.9
SCLOW	17.6	21.4	19.9	22.2	21.2	17.0	15.9	13.5	16.6	14.2	14.7	13.5	13.6
INDEP	21.6	16.6	16.3	13.2	20.6	21.9	31.7	31.8	25.5	34.8	31.5	32.5	19.9
CURR0	5.4	4.4	3.4	13.3	15.2	13.4	4.3	5.7	9.6	6.5	4.6	4.7	32.5
CURR1	46.8	52.3	38.8	41.4	44.5	13.5	6.8	18.7	12.2	19.9	27.1	39.1	15.4
CURR2	40.0	38.5	41.9	33.3	34.2	37.6	39.0	45.2	46.6	52.0	50.4	43.4	38.6
CURR3	7.8	4.8	15.9	12.0	6.2	35.6	49.9	30.4	31.5	21.7	17.9	12.8	13.5
JOINTDEG	8.6	11.2	17.6	19.7	8.4	16.2	35.2	7.2	31.0	4.0	25.0	12.9	65.8
ALMATH	47.7	84.7	96.4	77.2	81.3	51.5	61.7	36.1	15.6	15.7	12.3	15.4	17.0
ALGENS	21.9	24.9	23.3	19.3	19.8	18.4	23.1	22.7	20.1	20.3	21.7	22.9	11.3
SCOREA	14.7	30.9	38.6	17.4	22.1	17.5	26.9	40.2	11.7	19.8	24.7	19.5	3.4
SCORED	44.5	32.4	25.0	38.1	39.9	20.4	14.6	7.4	30.9	28.1	20.8	29.4	68.1
ALCOUNT	3.2	3.5	3.6	3.4	3.4	3.3	3.3	3.4	3.2	3.2	3.3	3.2	2.9
HCOUNT	5.1	5.1	5.0	4.9	4.9	5.7	5.7	5.8	4.9	4.8	4.7	5.0	3.9

Table 2.8 Subject-specific marginal effects: degree performance and socio-economic variables

		PP(mean)	DFIRST	POORDEG	SCLOW	INDEP
BIOL	WORK	44.2	-16.2*	5.8*	0.9	3.2*
(R^2=5.4)	OLFU	17.5	-6.4*	11.7*	2.4*	0.0
PHYS	WORK	32.0	-7.9*	15.0*	2.5*	2.7
(R^2=7.1)	OLFU	13.2	-1.2*	10.7*	1.5	-1.0
MATHS	WORK	49.3	-2.1*	-2.3	0.2	3.1
(R^2=5.3)	OLFU	14.7	-8.2*	8.0*	0.5	-0.7
COMP	WORK	68.3	-2.9*	-2.3*	-2.6	0.2
(R^2=8.9)	OLFU	15.6	-9.8*	10.6*	1.7	-0.1
ENGIN	WORK	64.6	-1.0*	-4.3*	-0.4	1.7
(R^2=5.9)	OLFU	14.4	-6.8*	10.7*	2.1	-1.2
BUS	WORK	75.0	2.8*	-3.2*	0.9	1.4
(R^2=8.9)	OLFU	12.9	-8.8*	6.2*	2.2	-1.4
ECON	WORK	62.8	-9.2*	-2.5*	-0.1	2.7
(R^2=5.5)	OLFU	17.1	-4.0*	9.0*	3.1	-3.7*
LAW	WORK	15.1	-5.7*	9.7*	0.3	-2.7*
(R^2=6.1)	OLFU	7.6	3.8*	6.7*	4.3*	0.1
OSOST [(a)]	WORK	51.2	11.0*	4.7*	-0.8	1.5
(R^2=2.8)	OLFU	18.7	-5.1*	5.3*	1.2	0.5
HUM	WORK	44.9	-13.4*	8.3*	1.3	1.0
(R^2=3.1)	OLFU	19.2	-11.8*	5.7*	2.0	-2.0*
CLAS	WORK	45.1	-9.9*	6.0*	0.2	2.3
(R^2=3.5)	OLFU	18.9	-1.3*	5.4*	0.3	-0.9
MEUL	WORK	55.2	-16.0*	1.7*	-0.6	0.2
(R^2=3.4)	OLFU	12.8	-6.3*	4.4*	2.9	0.5
EDU	WORK	87.1	1.8	1.7*	1.1	1.8
(R^2=22.7)	OLFU	8.6	-0.2*	0.6*	0.3	-0.7

* denotes statistical significance at 5% level

[(a)] OSOST=POL+SOCIO+OSOSCI

Table 2.9 Correlation coefficients (with p-values) between indicators of 'risk' and 'return'

	RET1[(a)]	RET2[(b)]	RISK1[(c)]	RISK2[(d)]	RISK3[(e)]
RET1	1				
RET2	0.208 (0.49)	1			
RISK1	0.841 (0.00)	0.121 (0.69)	1		
RISK2	-0.26 (0.38)	-0.121 (0.69)	0.126 (0.69)	1	
RISK3	0.753 (0.003)	0.088 (0.78)	0.940 (0.00)	0.104 (0.73)	1

(a) $RET1_j=P(WORK)_j$.

(b) $RET2_j=1-P(OLFU)$.

(c) $RISK1_j$=marginal effect of POORDEG on $P(WORK)_j$.

(d) $RISK2_j$=marginal effect of POORDEG on $P(OLFU)_j$.

(e) $RISK3_j = RISK1_j * \frac{1}{n_j}\sum_i^{n_j} POORDEG_i$.

Table 2.10 Subject-specific marginal effects: curriculum variables

		PP(mean)	CURR2	CURR3	ALGENS	ALMATH
BIOL	WORK	44.2	2.9*	3.7	4.6*	1.2
(R^2=5.4)	OLFU	17.5	0.0	0.4	-1.0	-1.3
PHYS	WORK	32.0	5.0	8.7	8.9*	2.0
(R^2=7.1)	OLFU	13.2	1.4	2.1	0.8	1.8
MATHS	WORK	49.3	5.4	0.6	0.4	-6.9
(R^2=5.3)	OLFU	14.7	0.3	-1.0	-0.2	-7.5
COMP	WORK	68.3	3.7*	7.8*	1.3	4.5
(R^2=8.9)	OLFU	15.6	-0.1	-2.9	-2.0	-0.4
ENGIN	WORK	64.6	4.6*	6.3*	2.1	-0.9
(R^2=5.9)	OLFU	14.4	-1.5	-1.4	-3.0	1.2
BUS	WORK	75.0	0.2	0.4	4.2*	1.3
(R^2=8.9)	OLFU	12.9	-0.5	-1.0	-0.1	-0.8
ECON	WORK	62.8	0.8	0.2	8.4*	8.3*
(R^2=5.5)	OLFU	17.1	-1.0	0.2	-3.2	-3.2
LAW	WORK	15.1	0.3	0.4	0.7	2.9*
(R^2=6.1)	OLFU	7.6	-0.9	0.7	1.3	0.7
OSOST	WORK	51.2	1.2	2.0	5.0*	1.8
(R^2=2.8)	OLFU	18.7	-0.1	0.1	-1.8	-2.6
HUM	WORK	44.9	2.0	1.0	2.1	5.0*
(R^2=3.1)	OLFU	19.2	1.7	0.3	-1.6	-2.6
CLAS	WORK	45.1	4.6*	5.5*	2.9	3.6
(R^2=3.5)	OLFU	18.9	-0.6	0.5	0.1	-0.5
MEUL	WORK	55.2	3.1	2.7	1.3	7.9*
(R^2=3.4)	OLFU	12.8	0.4	-0.1	1.2	-1.6
EDU	WORK	87.1	0.3	0.4	3.1	3.8
(R^2=22.7)	OLFU	8.6	0.2	-1.1	-0.5	-4.8*

* denotes statistical significance at 5% level.

Appendix 2A Variables definition ([a])

variable groups	Notes
Subject studied [(1)]: BIOL (Biology), CHEM (Chemistry), PHYS (Physics), OBIOL (Other Life Sciences), OPHYS (Other Physical Sciences), MATHS (Mathematical Sciences), COMP (Computer Science), ECON (Economics), SOCIO (Sociology), POL (Politics), LAW (Law), OSOSCI (Other Social Sciences), CLAS (Classics), ART (Creative Arts), ALMED (Allied Medicine), MEUL (Modern European Languages), HUM (Humanities), BUS (Business), ENGI (Engineering), EDU (Education Studies), OTHER (other minor courses). JOINTDEG[(2)] (='1' if student took a joint or combined degree course, '0' otherwise)	[(1)] Ex: BIOL='1' if student graduated in Biology, '0' otherwise. [(2)] This variable is different from the other subject dummies, in that it 'cuts across' all subject groups.
Type of institution attended [(2)]: OCIV ('old civic', like Belfast, Birmingham, Durham, Cardiff, Manchester, London, Liverpool), NCIV ('new civic', like Exeter, Newcastle, Nottingham, Southampton), OXBR (Oxford and Cambridge), EXCAT (Ex Colleges of Advanced Technology, like Brunel, Aston, Heriott-Watt), NEW60 (universities established in the 1960s, like East Anglia, Essex, Keele, Warwick), OSCOT (Other Scottish universities, like Glasgow, Edinburgh), OWEL (Other Welsh universities, like Aberystwyth, Bangor).	[(2)] This classification is the same adopted by the USR.
Degree class: DFIRST (first honours), DUSEC (upper second honours), DLSEC (lower second honours), DTHIRD (third honours or lower)[(3)]	[(3)] This group includes third, fourth, pass, and unclassified class honours.
Personal characteristics: MALE ('1' if student is male, '0' otherwise), MARRIED ('1' if student is married, '0' otherwise), MATURE ('1' is student is aged 21 or older at the date of enrolment, '0' otherwise)	

([a]) Unless specified otherwise, variables are categorical (i.e. categories are mutually exclusive).

Appendix 2A (continued)

variable groups	Notes
Social class: SC I (parent has a professional occupation), SC II (intermediate)[4], SC IIINM (skilled non-manual), SC IIIM (skilled manual), SC IV-V (partly skilled & unskilled), SC OTH (other)[5].	[4]Managerial and technical occupations [5]SC OTH includes individuals whose parental occupation is either inadequately described or unknown)
A-level point score [6] : SCOREA (28 pts or higher), SCOREB (23-27 pts), SCOREC (19-23 pts), SCORED (18 pts or less) *A-level curriculum* CURR0 (no A-level qualifications) CURR1 (one subject area of specialisation), CURR2 (two subject areas), CURR3 (three subject areas) [7]. ALGENS (=1 if student took A-level General Studies, 0 otherwise), ALMATH (=1 if student took A-level Mathematics, 0 otherwise) ALCOUNT (total number of A-level passes), HCOUNT (total number of Higher passes) [8] *School type*: LEA (comprehensive, grammar, special schools), INDEP (independent), NOSCH (school type information unavailable or unknown) *Other entry qualifications*: OENTQ0 (A-level), OENTQ1 (Vocational qualifications, like BTEC HNDs or HNCs, GNVQs), OENTQ2 (no formal qualifications and other minor qualifications like overseas diplomas).	[6] The 'point score' is an aggregated measure of an individual's A-level grades, calculated as follows: A=10, B=8, C=6, D=4, E=2 giving a possible maximum score of 30 points in the best three GCE A-level passes, and A=3, B=2, C=1 giving a possible maximum score of 15 points in the best five SCE Higher passes. The conversion Higher-to-A-level point score to create unique bands of grades is: A (14+→28+), B (12-13→23-27), C (10-11→19-23), D (9-→18-). [7] The A-level (Higher) subjects defining the subject areas are: *Science* (Biology, Chemistry, Physics, Other Sciences, Design, Electronics, Mechanics, Computers, Mathematics, Statistics); *Social Sciences* (Economics, Politics, Law); *Humanities* (Classics, English, Geography, History, French, German, Italian, Other Languages); *Other* (Art, Business, General Studies, Other) [8] ALCOUNT and HCOUNT are treated as 'continuous' variables.
Other course characteristics PTIME ('1' if method of study is part-time, '0' otherwise), CDUR ('1' if course duration >3 years, '0' otherwise	

Appendix 2A (continued)

variable groups	Notes
Residence prior to entry: RESID0-RESID10 (Non-resident UK nationals, Scotland and Northern Ireland, Wales, London, North-West, North & Yorkshire, West Midlands, East Midlands, South-West, South-East, East Anglia)	
Department-level information: DPRAE1 (RAE score ≤2), DPRAE2 (RAE score ≤4), DPRAE3 (RAE score>4), DPMALE (proportion of male leavers in student's department), DPSSR (staff/student ratio), DPPG (proportion of postgraduates), DP211 (proportion of leavers with first or upper second class honours)[10]	[10] DPMALE, DPSSR, DPPG, and DP211 are treated as 'continuous' variables.

Appendix 2B Secondary results

Table B.2.1 Multinomial logit estimates: all students (TRAIN is default)

	PWORK		TWORK		PSTUDY		UN		OLF	
variable	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
BIOL	-0.493	*0.096*	-0.051	*0.137*	1.170	*0.107*	0.029	*0.116*	0.234	*0.140*
OBIOL	-0.307	*0.099*	0.222	*0.140*	0.608	*0.115*	0.045	*0.123*	0.281	*0.146*
CHEM	-0.849	*0.102*	-0.719	*0.156*	1.167	*0.109*	-0.437	*0.122*	-0.431	*0.162*
PHYS	-1.078	*0.107*	-0.699	*0.159*	0.680	*0.115*	-0.384	*0.125*	-0.802	*0.171*
OPHYS	-0.469	*0.098*	0.090	*0.135*	0.531	*0.111*	-0.094	*0.118*	0.154	*0.140*
MATHS	-1.038	*0.086*	-1.015	*0.136*	-0.389	*0.104*	-0.790	*0.108*	-0.842	*0.141*
COMP	0.841	*0.144*	0.219	*0.194*	0.602	*0.160*	0.761	*0.160*	0.193	*0.212*
SOCIO	-0.462	*0.106*	0.081	*0.153*	-0.087	*0.133*	0.040	*0.131*	0.054	*0.163*
POL	-0.771	*0.097*	-0.323	*0.139*	-0.261	*0.117*	-0.359	*0.119*	-0.548	*0.154*
LAW	-3.298	*0.088*	-2.807	*0.154*	-2.328	*0.111*	-2.941	*0.127*	-2.078	*0.134*
OSOSCI	-0.615	*0.095*	-0.190	*0.137*	-0.054	*0.116*	-0.487	*0.122*	-0.178	*0.144*
CLAS	-1.096	*0.091*	-0.534	*0.134*	-0.512	*0.111*	-0.415	*0.114*	-0.567	*0.139*
ART	-0.989	*0.108*	-0.829	*0.173*	-0.514	*0.138*	-0.653	*0.143*	-0.634	*0.177*
ALMED	-1.296	*0.101*	-1.036	*0.157*	-1.120	*0.132*	-0.889	*0.132*	-1.113	*0.167*
MEUL	1.299	*0.140*	0.256	*0.203*	1.550	*0.158*	0.467	*0.173*	0.412	*0.213*
ENGIN	0.394	*0.094*	0.096	*0.130*	0.488	*0.108*	0.240	*0.110*	-0.031	*0.140*
BUS	0.304	*0.093*	0.213	*0.132*	-0.487	*0.122*	0.009	*0.116*	0.249	*0.141*
HUM	-1.059	*0.082*	-0.601	*0.120*	-0.502	*0.099*	-0.489	*0.100*	-0.702	*0.124*
EDU	0.428	*0.156*	-0.006	*0.235*	-0.755	*0.238*	-0.238	*0.209*	0.241	*0.248*
OTHER	-0.047	*0.093*	-0.137	*0.138*	0.230	*0.111*	-0.096	*0.115*	0.058	*0.143*
OXBR	-0.114	*0.063*	-0.650	*0.113*	-0.290	*0.075*	-0.629	*0.088*	-0.706	*0.100*
NCIV	0.075	*0.037*	-0.250	*0.060*	-0.186	*0.045*	0.029	*0.047*	-0.111	*0.060*
EXCAT	0.315	*0.051*	0.293	*0.074*	-0.132	*0.061*	0.100	*0.063*	-0.180	*0.085*
NEW60	0.051	*0.039*	-0.108	*0.060*	-0.151	*0.047*	-0.084	*0.050*	-0.252	*0.066*
OSCOT+OWEL	-0.319	*0.051*	-0.003	*0.074*	-0.061	*0.060*	-0.134	*0.064*	-0.264	*0.086*
DFIRST	0.256	*0.050*	-0.348	*0.101*	1.040	*0.053*	-0.321	*0.078*	0.093	*0.084*
DLSEC	0.010	*0.030*	0.403	*0.045*	-0.739	*0.038*	0.484	*0.038*	0.197	*0.049*
DTHIRD	-0.176	*0.052*	0.423	*0.074*	-1.573	*0.071*	0.655	*0.060*	0.111	*0.086*
SC I	-0.119	*0.035*	-0.069	*0.056*	0.070	*0.042*	-0.100	*0.046*	-0.007	*0.056*
SC IIINM	0.017	*0.043*	0.130	*0.065*	0.003	*0.051*	0.026	*0.055*	-0.047	*0.071*
SC IIIM	-0.081	*0.045*	0.077	*0.068*	-0.075	*0.053*	0.060	*0.056*	-0.231	*0.080*
SC IV-V	0.044	*0.055*	0.332	*0.079*	0.071	*0.065*	0.323	*0.066*	-0.098	*0.096*
SC OTH	-0.074	*0.054*	0.074	*0.082*	0.033	*0.064*	0.192	*0.066*	-0.026	*0.088*
NOSCH	-0.045	*0.109*	-0.123	*0.171*	0.283	*0.123*	-0.013	*0.136*	0.233	*0.166*
INDEP	0.009	*0.031*	-0.165	*0.050*	-0.122	*0.038*	-0.273	*0.042*	0.104	*0.051*
RES0	-0.419	*0.086*	-0.470	*0.146*	0.490	*0.096*	-0.296	*0.114*	-0.058	*0.127*
RES1	-0.998	*0.060*	-0.780	*0.090*	-0.435	*0.072*	-0.869	*0.078*	-1.095	*0.105*
RES2	-0.475	*0.064*	-0.484	*0.102*	-0.174	*0.078*	-0.361	*0.083*	-0.495	*0.109*
RES3	-0.154	*0.055*	-0.125	*0.083*	0.143	*0.065*	-0.003	*0.069*	-0.181	*0.086*
RES4	-0.318	*0.052*	-0.385	*0.081*	-0.184	*0.063*	-0.165	*0.066*	-0.473	*0.086*
RES5	-0.243	*0.054*	-0.298	*0.084*	-0.061	*0.065*	-0.051	*0.068*	-0.394	*0.089*
RES6	-0.202	*0.058*	-0.282	*0.090*	-0.100	*0.070*	-0.139	*0.074*	-0.363	*0.094*
RES7	-0.254	*0.065*	-0.148	*0.097*	-0.162	*0.078*	-0.173	*0.083*	-0.309	*0.106*
RES8	-0.050	*0.055*	-0.013	*0.082*	-0.126	*0.067*	0.009	*0.070*	-0.055	*0.084*
RES10	-0.197	*0.055*	-0.203	*0.084*	-0.095	*0.066*	-0.136	*0.071*	-0.167	*0.085*
OENTQ1	0.362	*0.097*	-0.032	*0.147*	0.373	*0.113*	0.240	*0.116*	-0.074	*0.160*
OENTQ2	0.329	*0.125*	0.058	*0.196*	0.131	*0.146*	0.110	*0.161*	-0.230	*0.201*
SCOREC	0.034	*0.038*	-0.040	*0.055*	-0.222	*0.045*	-0.011	*0.047*	-0.003	*0.063*
SCOREB	0.191	*0.040*	-0.057	*0.061*	-0.224	*0.048*	-0.007	*0.051*	0.123	*0.066*
SCOREA	0.306	*0.046*	-0.105	*0.075*	-0.130	*0.055*	0.013	*0.061*	0.319	*0.076*
ALMATH	0.240	*0.036*	0.080	*0.054*	0.204	*0.042*	0.101	*0.045*	0.119	*0.056*
ALCOUNT	-0.003	*0.016*	0.001	*0.024*	0.073	*0.018*	0.023	*0.020*	-0.065	*0.027*
HCOUNT	0.033	*0.014*	0.017	*0.022*	0.048	*0.017*	0.075	*0.018*	0.057	*0.025*
MARRIED	0.014	*0.140*	-0.470	*0.251*	-0.282	*0.161*	-0.608	*0.184*	-0.646	*0.302*
FEMALE	-0.295	*0.029*	-0.329	*0.044*	-0.605	*0.034*	-0.846	*0.038*	-0.251	*0.046*
FEM*MAR	-0.080	*0.164*	0.110	*0.307*	0.331	*0.194*	0.664	*0.220*	1.420	*0.324*
CDUR	0.593	*0.040*	0.248	*0.060*	0.253	*0.048*	0.227	*0.051*	0.268	*0.065*
PTIME	0.771	*0.129*	-0.490	*0.265*	0.261	*0.159*	-0.141	*0.179*	-0.193	*0.256*
MATURE	0.152	*0.051*	-0.087	*0.082*	0.256	*0.061*	0.246	*0.064*	-0.029	*0.089*
DPRAE2	0.206	*0.031*	0.102	*0.047*	0.105	*0.037*	0.122	*0.039*	0.205	*0.051*
DPRAE3	0.246	*0.040*	0.071	*0.061*	0.204	*0.047*	0.201	*0.050*	0.420	*0.063*
DPR211	-0.282	*0.099*	0.584	*0.154*	-0.401	*0.123*	0.400	*0.128*	0.928	*0.167*
DPMALE	0.289	*0.118*	0.575	*0.180*	1.107	*0.144*	0.594	*0.152*	0.778	*0.190*
DPPG	-0.242	*0.147*	-0.253	*0.228*	0.293	*0.172*	-0.049	*0.186*	-0.305	*0.238*
constant	1.730	*0.143*	-0.695	*0.217*	0.012	*0.172*	-0.181	*0.181*	-1.085	*0.230*
N	63706									
LL	-86458.2									
pseudo R2	0.11									

Table B.2.2 Multinomial logit estimates: females (TRAIN is default)

	PWORK		TWORK		PSTUDY		UN		OLF	
variable	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
BIOL	-0.639	*0.140*	0.003	*0.218*	1.138	*0.169*	-0.163	*0.185*	0.139	*0.213*
OBIOL	-0.355	*0.143*	0.281	*0.219*	0.609	*0.179*	-0.189	*0.192*	0.095	*0.221*
CHEM	-1.125	*0.154*	-0.761	*0.262*	0.969	*0.177*	-0.948	*0.217*	-0.603	*0.256*
PHYS	-1.298	*0.205*	-0.820	*0.360*	0.748	*0.220*	-0.398	*0.255*	-1.384	*0.402*
OPHYS	-0.759	*0.149*	-0.061	*0.229*	0.392	*0.180*	-0.476	*0.199*	-0.182	*0.230*
MATHS	-1.231	*0.137*	-1.069	*0.241*	-0.452	*0.176*	-0.947	*0.191*	-1.090	*0.235*
COMP	0.098	*0.260*	0.090	*0.403*	0.133	*0.315*	0.063	*0.330*	-0.297	*0.451*
SOCIO	-0.427	*0.150*	0.105	*0.232*	0.044	*0.196*	-0.209	*0.201*	-0.024	*0.237*
POL	-0.760	*0.154*	-0.012	*0.231*	-0.124	*0.195*	-0.539	*0.205*	-0.321	*0.241*
LAW	-3.230	*0.136*	-2.742	*0.251*	-2.071	*0.178*	-2.887	*0.203*	-1.998	*0.211*
OSOSCI	-0.790	*0.141*	-0.154	*0.219*	-0.102	*0.182*	-0.720	*0.195*	-0.231	*0.220*
CLAS	-1.170	*0.136*	-0.469	*0.216*	-0.520	*0.177*	-0.779	*0.184*	-0.578	*0.213*
ART	-1.072	*0.154*	-0.684	*0.251*	-0.513	*0.207*	-0.930	*0.217*	-0.669	*0.249*
ALMED	-1.434	*0.146*	-0.965	*0.237*	-1.115	*0.198*	-1.228	*0.204*	-1.239	*0.242*
MEUL	1.183	*0.182*	0.307	*0.277*	1.456	*0.219*	0.286	*0.240*	0.169	*0.292*
ENGIN	0.064	*0.173*	0.175	*0.262*	0.166	*0.213*	-0.224	*0.229*	-0.067	*0.267*
BUS	0.193	*0.141*	0.303	*0.219*	-0.393	*0.195*	-0.225	*0.193*	0.190	*0.222*
HUM	-1.121	*0.129*	-0.514	*0.205*	-0.541	*0.166*	-0.759	*0.172*	-0.680	*0.200*
EDU	0.519	*0.211*	0.251	*0.313*	-0.585	*0.321*	-0.270	*0.287*	0.222	*0.341*
OTHER	-0.369	*0.142*	-0.180	*0.231*	0.250	*0.181*	-0.481	*0.195*	-0.266	*0.234*
OXBR	0.050	*0.088*	-0.271	*0.155*	-0.014	*0.111*	-0.525	*0.140*	-0.461	*0.143*
NCIV	0.081	*0.049*	-0.193	*0.084*	-0.164	*0.066*	0.069	*0.071*	-0.096	*0.083*
EXCAT	0.280	*0.068*	0.265	*0.107*	-0.206	*0.091*	0.215	*0.097*	-0.135	*0.117*
NEW60	0.087	*0.051*	-0.066	*0.080*	-0.131	*0.067*	-0.060	*0.075*	-0.302	*0.090*
OSCOT+OWEL	-0.348	*0.066*	-0.076	*0.103*	-0.105	*0.084*	-0.052	*0.092*	-0.264	*0.111*
DFIRST	0.244	*0.073*	-0.282	*0.147*	1.109	*0.078*	-0.055	*0.121*	0.220	*0.119*
DLSEC	-0.020	*0.040*	0.332	*0.061*	-0.766	*0.056*	0.381	*0.057*	0.095	*0.067*
DTHIRD	-0.049	*0.076*	0.275	*0.125*	-1.666	*0.130*	0.514	*0.104*	0.301	*0.129*
SC I	-0.075	*0.047*	-0.005	*0.078*	0.098	*0.059*	-0.043	*0.070*	-0.006	*0.077*
SC IIINM	0.001	*0.056*	0.023	*0.093*	0.016	*0.074*	-0.038	*0.084*	-0.072	*0.098*
SC IIIM	-0.051	*0.059*	0.087	*0.095*	-0.078	*0.079*	-0.014	*0.087*	-0.239	*0.111*
SC IV-V	0.162	*0.076*	0.560	*0.107*	0.096	*0.098*	0.398	*0.101*	0.112	*0.128*
SC OTH	-0.099	*0.071*	0.151	*0.113*	-0.017	*0.093*	0.225	*0.098*	-0.103	*0.119*
NOSCH	0.062	*0.143*	0.070	*0.228*	0.198	*0.177*	0.138	*0.199*	0.101	*0.207*
INDEP	0.003	*0.042*	-0.133	*0.071*	-0.041	*0.056*	-0.213	*0.065*	0.155	*0.070*
RES0	-0.340	*0.113*	-0.458	*0.200*	0.394	*0.134*	-0.195	*0.167*	-0.163	*0.174*
RES1	-0.983	*0.078*	-0.525	*0.118*	-0.430	*0.102*	-0.776	*0.118*	-1.061	*0.141*
RES2	-0.518	*0.084*	-0.434	*0.139*	-0.246	*0.112*	-0.530	*0.132*	-0.496	*0.144*
RES3	-0.092	*0.073*	-0.239	*0.120*	0.121	*0.092*	0.014	*0.105*	-0.211	*0.116*
RES4	-0.305	*0.070*	-0.465	*0.117*	-0.241	*0.092*	-0.122	*0.102*	-0.472	*0.117*
RES5	-0.183	*0.073*	-0.173	*0.117*	-0.035	*0.094*	-0.025	*0.106*	-0.407	*0.123*
RES6	-0.196	*0.077*	-0.240	*0.126*	-0.204	*0.102*	-0.077	*0.113*	-0.479	*0.132*
RES7	-0.274	*0.085*	-0.250	*0.138*	-0.261	*0.113*	-0.164	*0.126*	-0.447	*0.148*
RES8	-0.111	*0.072*	-0.116	*0.116*	-0.206	*0.096*	0.068	*0.104*	-0.170	*0.114*
RES10	-0.313	*0.073*	-0.253	*0.117*	-0.232	*0.094*	-0.139	*0.107*	-0.272	*0.115*
OENTQ1	0.187	*0.137*	0.055	*0.215*	0.439	*0.172*	0.211	*0.177*	0.100	*0.213*
OENTQ2	0.334	*0.167*	0.120	*0.267*	0.481	*0.210*	0.159	*0.239*	0.063	*0.251*
SCOREC	0.073	*0.048*	0.037	*0.075*	-0.186	*0.063*	0.034	*0.069*	0.055	*0.084*
SCOREB	0.167	*0.051*	-0.016	*0.083*	-0.295	*0.068*	-0.019	*0.076*	0.169	*0.088*
SCOREA	0.328	*0.062*	-0.060	*0.106*	-0.142	*0.079*	0.042	*0.097*	0.340	*0.104*
ALMATH	0.246	*0.048*	0.074	*0.075*	0.168	*0.060*	-0.087	*0.071*	0.154	*0.079*
ALCOUNT	-0.006	*0.021*	0.031	*0.034*	0.110	*0.027*	0.008	*0.031*	-0.065	*0.037*
HCOUNT	0.045	*0.019*	0.032	*0.029*	0.090	*0.024*	0.092	*0.027*	0.046	*0.034*
PTIME	0.674	*0.174*	-0.391	*0.359*	0.248	*0.223*	-0.428	*0.277*	-0.359	*0.338*
MATURE	0.281	*0.073*	0.149	*0.121*	0.525	*0.090*	0.398	*0.100*	0.328	*0.122*
MARRIED	-0.190	*0.110*	-0.646	*0.209*	-0.140	*0.135*	-0.155	*0.155*	0.482	*0.164*
CDUR	0.632	*0.051*	0.286	*0.083*	0.288	*0.068*	0.311	*0.074*	0.303	*0.087*
DPRAE2	0.227	*0.041*	0.085	*0.065*	0.119	*0.053*	0.091	*0.059*	0.155	*0.069*
DPRAE3	0.233	*0.053*	0.014	*0.088*	0.220	*0.067*	0.296	*0.074*	0.374	*0.085*
DPR211	-0.308	*0.128*	0.719	*0.212*	-0.362	*0.176*	0.352	*0.190*	0.745	*0.228*
DPMALE	0.201	*0.156*	0.288	*0.244*	1.331	*0.203*	0.681	*0.224*	0.573	*0.265*
DPPG	-0.352	*0.200*	0.000	*0.325*	0.068	*0.251*	0.069	*0.288*	-0.667	*0.341*
constant	1.552	*0.197*	-1.300	*0.316*	-0.835	*0.257*	-0.830	*0.278*	-1.039	*0.332*
N	29391									
LL	-40093.5									
pseudo R2	0.09									

Table B.2.3 Multinomial logit estimates: males (TRAIN is default)

	PWORK		TWORK		PSTUDY		UN		OLF	
variable	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
BIOL	-0.358	0.145	0.059	0.196	1.257	0.153	0.143	0.164	0.340	0.202
OBIOL	-0.417	0.153	0.212	0.210	0.581	0.168	0.083	0.177	0.477	0.212
CHEM	-0.610	0.141	-0.564	0.203	1.372	0.147	-0.131	0.159	-0.244	0.214
PHYS	-0.994	0.130	-0.763	0.186	0.713	0.138	-0.336	0.148	-0.637	0.198
OPHYS	-0.198	0.138	0.306	0.178	0.729	0.150	0.185	0.156	0.472	0.185
MATHS	-0.912	0.113	-0.949	0.170	-0.313	0.132	-0.706	0.135	-0.648	0.179
COMP	1.092	0.177	0.240	0.231	0.823	0.193	1.011	0.193	0.416	0.250
SOCIO	-0.765	0.171	0.159	0.234	-0.346	0.204	0.029	0.194	0.067	0.257
POL	-0.810	0.127	-0.484	0.180	-0.357	0.149	-0.323	0.148	-0.756	0.204
LAW	-3.446	0.120	-2.794	0.202	-2.569	0.146	-3.098	0.166	-2.217	0.182
OSOSCI	-0.429	0.141	-0.106	0.196	0.057	0.164	-0.367	0.169	-0.146	0.209
CLAS	-1.129	0.133	-0.449	0.195	-0.497	0.157	-0.163	0.156	-0.749	0.206
ART	-1.012	0.168	-0.989	0.288	-0.526	0.200	-0.549	0.206	-0.720	0.286
ALMED	-1.222	0.159	-1.008	0.256	-1.085	0.200	-0.642	0.194	-1.070	0.269
MEUL	1.306	0.261	0.173	0.365	1.717	0.279	0.449	0.301	0.719	0.357
ENGIN	0.526	0.116	0.017	0.159	0.629	0.131	0.411	0.133	0.065	0.171
BUS	0.348	0.127	0.187	0.174	-0.555	0.161	0.076	0.150	0.280	0.188
HUM	-1.044	0.108	-0.570	0.156	-0.465	0.126	-0.374	0.127	-0.753	0.165
EDU	-0.196	0.236	-0.993	0.460	-1.212	0.370	-0.920	0.331	0.142	0.371
OTHER	0.263	0.133	0.039	0.186	0.371	0.152	0.227	0.154	0.396	0.191
OXBR	-0.280	0.093	-1.056	0.166	-0.510	0.103	-0.752	0.118	-0.921	0.143
NCIV	0.060	0.058	-0.306	0.087	-0.202	0.066	-0.004	0.067	-0.118	0.090
EXCAT	0.385	0.077	0.368	0.105	-0.036	0.087	0.116	0.088	-0.190	0.122
NEW60	0.008	0.061	-0.160	0.090	-0.168	0.069	-0.108	0.071	-0.188	0.099
OSCOT+OWEL	-0.276	0.082	0.113	0.111	-0.001	0.090	-0.141	0.094	-0.275	0.134
DFIRST	0.280	0.071	-0.362	0.139	1.013	0.073	-0.444	0.103	-0.001	0.120
DLSEC	0.056	0.048	0.488	0.068	-0.702	0.055	0.578	0.055	0.295	0.073
DTHIRD	-0.257	0.071	0.470	0.098	-1.586	0.090	0.673	0.078	-0.032	0.116
SC I	-0.180	0.054	-0.153	0.082	0.023	0.060	-0.162	0.065	-0.019	0.081
SC IIINM	0.048	0.067	0.234	0.093	0.018	0.075	0.080	0.078	0.002	0.104
SC IIIM	-0.116	0.069	0.062	0.098	-0.089	0.077	0.076	0.078	-0.230	0.115
SC IV-V	-0.119	0.081	0.059	0.116	-0.037	0.090	0.190	0.091	-0.363	0.144
SC OTH	-0.052	0.084	-0.019	0.121	0.062	0.094	0.174	0.095	0.037	0.134
NOSCH	-0.161	0.165	-0.372	0.255	0.294	0.177	-0.160	0.190	0.384	0.264
INDEP	0.019	0.047	-0.190	0.071	-0.177	0.055	-0.316	0.057	0.064	0.074
RES0	-0.509	0.133	-0.480	0.216	0.525	0.140	-0.390	0.161	0.029	0.189
RES1	-1.011	0.093	-1.060	0.141	-0.439	0.105	-0.929	0.110	-1.131	0.160
RES2	-0.385	0.102	-0.484	0.153	-0.060	0.115	-0.207	0.117	-0.485	0.168
RES3	-0.217	0.084	-0.037	0.119	0.137	0.094	-0.033	0.097	-0.158	0.127
RES4	-0.318	0.079	-0.323	0.114	-0.142	0.090	-0.171	0.092	-0.466	0.126
RES5	-0.318	0.082	-0.446	0.123	-0.116	0.092	-0.105	0.094	-0.400	0.129
RES6	-0.204	0.090	-0.324	0.131	-0.038	0.101	-0.170	0.104	-0.245	0.137
RES7	-0.190	0.100	-0.012	0.141	-0.051	0.113	-0.125	0.118	-0.137	0.154
RES8	0.048	0.086	0.124	0.119	-0.021	0.098	0.036	0.100	0.103	0.125
RES10	-0.029	0.086	-0.095	0.123	0.075	0.096	-0.041	0.101	-0.002	0.127
OENTQ1	0.459	0.140	-0.120	0.205	0.375	0.156	0.297	0.159	-0.311	0.245
OENTQ2	0.271	0.189	-0.033	0.288	-0.136	0.208	0.038	0.220	-0.565	0.324
SCOREC	-0.046	0.061	-0.153	0.083	-0.296	0.068	-0.087	0.068	-0.099	0.095
SCOREB	0.208	0.065	-0.092	0.092	-0.180	0.072	0.003	0.074	0.063	0.100
SCOREA	0.261	0.072	-0.146	0.109	-0.146	0.079	-0.029	0.085	0.265	0.112
ALMATH	0.264	0.054	0.091	0.079	0.246	0.061	0.200	0.064	0.083	0.081
ALCOUNT	0.000	0.023	-0.021	0.034	0.052	0.026	0.031	0.027	-0.064	0.039
HCOUNT	0.009	0.022	0.001	0.033	0.008	0.024	0.053	0.026	0.063	0.037
PTIME	0.872	0.197	-0.572	0.399	0.295	0.233	0.039	0.247	-0.009	0.394
MARRIED	0.192	0.156	-0.129	0.264	-0.069	0.175	-0.377	0.197	-0.309	0.320
CDUR	0.525	0.063	0.175	0.089	0.193	0.071	0.122	0.073	0.227	0.098
MATURE	0.028	0.073	-0.287	0.114	0.041	0.083	0.102	0.084	-0.418	0.133
DPRAE2	0.172	0.047	0.111	0.068	0.087	0.053	0.121	0.055	0.264	0.077
DPRAE3	0.257	0.061	0.150	0.087	0.204	0.068	0.159	0.070	0.484	0.093
DPR211	-0.242	0.158	0.440	0.230	-0.426	0.182	0.409	0.184	1.142	0.246
DPMALE	0.183	0.188	0.891	0.281	0.780	0.217	0.457	0.220	0.778	0.283
DPPG	-0.126	0.218	-0.484	0.326	0.481	0.245	-0.073	0.255	-0.060	0.337
constant	1.772	0.210	-0.579	0.312	0.287	0.241	-0.161	0.247	-1.239	0.321
N	34315									
LL	-46054.3									
pseudo R2	0.11									

Chapter 3

The determinants of graduates' first destinations and the business cycle in the UK: evidence from the USR, 1980-1993

3.1 Introduction

This chapter extends the first destination analysis presented in Chapter 2 to fourteen consecutive cohorts of graduates who left university between 1980 and 1993.[1] Introducing a dynamic dimension to the analysis is not just about replicating results for earlier waves of university students. The demand and supply of graduates depend on time-specific factors, which are directly related to the business cycle as well as to the changing institutional framework of the higher education system. This is particularly the case for newly qualified graduates. New entrants to the labour market can be particularly vulnerable to demand and supply shifts. Starting a career in an adverse economic climate can prove particularly difficult because recruitment is often the first casualty of a downswing in demand. Similarly, supply-side shocks like a surge in enrolment rates or changes in the private cost of a university degree can lead to short-term surpluses or shortages of highly educated labour. For instance, an increase in university tuition fees is expected to affect students' time preferences and job search patterns. New graduates may become more likely to take up temporary jobs or enter occupations for which they are overqualified to pay off their debts, especially when demand is slack. Alternatively, the start of a career can be delayed by a period of further study or training.

During the 1980s and early 1990s both sides of the graduate labour market underwent significant structural changes. On the supply side, the last two decades were characterised by an unprecedented expansion in university admissions fuelled by a series of reforms targeted at making higher education a mass system.

Figures collected from a variety of sources indicate a four-fold increase in the number of first-degree students between 1966 and 1996 (Dearing, 1997; Greenaway and Williams, 1973). In the period 1979-1996 undergraduate numbers, excluding the Open University, grew from 671 thousands to 1,481 thousands (Dearing, 1997, p. 19). On the demand side, the graduate labour market broadly mimicked the trends in the wider labour market characterised by two periods of significant contraction in the early 1980s and early 1990s and by a long spell of steady expansion in the mid and late 1980s.

Amid this dynamic context, it is interesting to look at how the effects of some of the key factors influencing graduates' first destinations discussed in Chapter 2 have changed over time and, in particular, how these effects relate to the business cycle. Supply and demand-side shifts in the graduate labour market can alter the impact of human capital inputs like university attended, subject studied, and degree class obtained on graduates' early careers. For instance, in periods of high unemployment the 'cream skimming' of graduates by employers based on those inputs is likely to become more widespread, in the sense that the employability *premium* of having graduated from a prestigious institution and/or having obtained a 'good' class of degree may be higher. As noted by Lynton (1993) in a OECD cross-country study '[...when participation was still relatively low university degrees tended to lend considerable status, high employability, and high returns with little differentiation among fields in which the degrees were awarded or among institutions. As a result of the rising participation rates into higher education, employers increasingly tend to attach a screening function to the

[1] Owing to major changes after 1980 in the USR coding procedure of some key first destination variables used in this study, we use only data starting from 1980 to ensure inter-cohort

most prestigious universities and to the most selective and demanding programmes]' (p. 6).

Second, socio-economic factors could equally be more influential when the labour market is slack. In Chapter 2 we found significant social class and independent school effects. One of the explanations offered there was the existence of family and social networks (Montgomery, 1991; Holzer, 1987). It may happen that these network effects on graduates' labour-market outcomes are counter-cyclical, in the sense of facilitating the start of a career particularly during economic downturns (Dolton, *et al.*, 1997).

Third, the pattern of gender effects may have changed over time in response to the changing role of women in society and to the growth of female labour market participation. For instance, the introduction of anti-discrimination legislation in the 1970s in the UK, aimed at improving the working condition of women, may have affected the early careers of female graduates vis-à-vis men (Dolton *et al.*, 1996).

Finally, the recent growth and diversification in the graduate population has posed new challenges for recruiters, universities, and new graduates alike, as they seek to meet and respond to each others' needs in an increasingly large, fragmented and diverse market place. One related concern is that the economy is generating university graduates at a faster rate than it is producing jobs for those graduates. New graduates may be increasingly unable to secure occupations that command the expected private rate of return to a university degree. This phenomenon is commonly referred to as *overeducation* and has recently received much attention in the UK literature (Dolton and Vignoles, 2000; Sloane *et al.*, 1999; Chevalier,

comparability.

2000). An interesting extension of the analysis presented in Chapter 2 is the distinction between occupations requiring a first degree ('graduate' jobs) and occupations for which a university degree is not required ('non-graduate' jobs). This distinction will enable us to address the issue of how personal and educational factors influence the probability of graduates to enter not simply an occupation, but jobs that adequately match up with their educational qualifications.

Broadly speaking, the extension of the analysis on graduates' first destinations to earlier cohorts of university leavers is expected to shed light on the graduate labour market trends in the UK during the period 1980-1993.

The rest of this chapter is organised as follows. Section 3.2 presents an outlook of the stylised facts of the graduate labour market in the UK in recent years. Section 3.3 briefly discusses the modelling strategy and clarifies some data issues. Summary statistics of some key variables are presented in Section 3.4, while in Section 3.5 we discuss the dynamics of the main factors that influence the first destination decisions of consecutive cohorts of graduates. Section 3.6 takes a closer look at the time profile of gender effects. Section 3.7 focuses on the distinction between 'graduate' and 'non-graduate' employment. Finally, Section 3.8 concludes with a summary of the main findings.

3.2 Higher education and the graduate labour market in the UK: the stylised facts

The 1980s and 1990s have seen remarkable changes both in the supply and demand of university graduates in the UK. Since the publication of the Robbins Report in 1963 higher education in the UK has rapidly evolved towards becoming a mass system. Undoubtedly, the abolition of the binary divide between 'Old Universities' and former Polytechnics in 1992 (Further and Higher Education Act), which brought the number of universities from around 80 to over 180, was by far the single most important factor behind the expansion in the supply of university graduates.[2] Alongside this crucial institutional reform, other factors may have directly or indirectly contributed to the growth of the higher education sector. First, the greater openness on the part of many institutions towards students with 'non-traditional' (other than A-level) entry backgrounds, including a number of vocational and technical routes into higher education like BTEC qualifications. Second, the comprehensive reorganisation of secondary education and the 1988 Education Reform Act. These reforms had the effect of removing the exclusionary perception of 'low ability' state-run schools[3] and non-academic track qualifications (O-levels as opposed to CSEs) as well as removing artificial

---

[2] Following the recommendations of the Robbins Report, ten colleges of advanced technology (CATs) were upgraded to technological universities in the early 1960s. In the same period, the establishment of seven completely new universities provided a remarkable expansion of university places (Greenaway and Williams, 1973).

[3] Until the late 1960s, secondary schooling in England and Wales was stratified into 'public' (i.e. private or independent) schools which cater for all ages of children, non-selective or comprehensive local education authority (LEA) schools for children aged 11 and above, and LEA selective schools which selected children at the age of 11 by an ability test, known as '11-plus'. State-run selective schools were divided into high-ability 'grammar' schools for children who passed the test, and low-ability 'secondary modern' schools for children who failed the '11-plus'. The reform removed the distinction between grammar and secondary modern schools, even though a very small number of LEAs still retain an 11 plus exam and operate grammar school alongside the comprehensive system (Harmon and Walker, 2000).

ceilings on the number of places available at the most popular schools.[4] Each of these measures are likely to have encouraged more pupils to stay on after 16 (Finegold *et al.*, 1992). Last, the accelerator effect of growing numbers of graduate parents (in particular the 1960s graduate boomers) on the incentive for children to go to university, in consideration of the fact that intergenerational aspirations tend to be iterative (Shuller, 1996). These social factors contributed to mitigate the diminishing size of the 18-21 year-olds cohort in the last two decades due to demographic stagnation. The marked expansion in the broader undergraduate student population is reflected in the growth of graduate output in the 1980s and early 1990s. Figure 3.1 shows the number of graduates from 'pre-1992' universities in the UK over the period 1980-1993. Between 1980 and 1993 the graduate population grew by 32%. It is also clear that the annual growth rate has been much higher in the early 1990s (over 4%) than in the 1980s (about 1%).[5] On the demand side, unemployment rates among newly qualified graduates mimicked closely the corresponding trend in the wider UK economy, as shown in Figure 3.2. Similar trends were also observed in more recent years by Elias (1999) using first destination and LFS data home-domiciled UK students over the period 1991-1999. Interestingly, Figure 3.2 shows that graduate unemployment

---

[4] Open enrolment, coupled with the introduction of technology colleges and grant-maintained schools has significantly increased parental choice. Furthermore, changes in formula funding based on age-weighted pupil numbers, the local management of schools and the opportunity to opt out of LEA control have significantly contributed to the creation of a quasi-market in the school's sector (Bradley *et al.*, 2000).

[5] When graduates from former Polytechnics are included into the count, the growth rate in graduate numbers over the period 1980-1993 exceeds 50%. The faster expansion of former Polytechnics sector is well-documented (Shuller, 1996; OECD, 1993) This is explained by the fact that in 1981 the government announced sharp cuts in higher education funding. The universities responded by cutting student intakes, while keeping expenditure per student stable. On the other hand, the Polytechnics steadily increased their intake, often attracting students who were diverted from universities (OECD, 1993, p. 65).

responded more quickly and dramatically than overall unemployment to changes in the business cycle.

The evidence provided in Figure 3.2 reflects some well-known facts. The early 1980s were for the British economy the turning point of long-lived economic recession, which started in the mid-1970s and hit its lowest point in 1981-1982. From 1983 the economic cycle inverted its course and a period of steady and rapid expansion followed for most of the rest of the decade. Between 1983 and 1988 British GDP grew at an average annual rate of 3.5% compared to 1.6% between 1970 and 1982 (OECD, 1993, p. 72). In those years, the Conservative government promoted competition and flexibility in the labour market. This was done through a tough deregulation policy targeted at removing the 'rigidities' such as job protection laws (unfair dismissal), legal immunities for Trade Unions, cuts in the value of welfare benefits and redundancy payments. Union density fell from 50% in the mid-1970s to 35% in the mid-1990s (LFS). The combination of a buoyant economy, a more flexible labour market that significantly increased the number of job opportunities, and non-increasing worked hours rapidly forced unemployment rates to fall back to the levels of the mid-1970s. By the end of the 1980s, the economic expansion came to an end and the British economy experienced a new downturn that reached its low in 1992. The 'bust' of the early 1990s has been widely attributed to financial fragility both at the corporate and household level. Financial liberalisation, greater competition among financial institutions and expectations of long-term income growth during the economic upturn of the 1980s contributed to the rise of corporate debt and leverage. As noted by Hall (2002), '[... the unexpected deterioration in economic prospects in the late 1980s may have led to a sharp downward revision to companies' desired levels of capital and debt]'. Concomitantly, the 20% rise in house prices over 1988/1989 alone in the UK and their ensuing fall in the early 1990s generated a risk of 'negative equity' positions for highly leveraged households (Davis, 1995). The burst of the 'asset-price' and 'house property' bubbles, led to a sharp cutback in spending and investment. In the early 1990s output growth plummeted from over 7% at the peak to -6% at the trough, and job-shedding was probably the more severe for the sudden change in fortunes. The business cycle for Great Britain is shown in Figure 3.3. From the diagram it is also evident that British manufacturing has been more sensitive to economic fluctuations than the services sector. In particular, the worst-affected industries in the early 1990s were the traditional manufacturing sectors of mechanical engineering, steel, transport equipment and textiles and clothing, together with those that suffered the knock-on effects of the recession in construction (wood products, mineral products).

3.3 Methodology

The modelling strategy used in this chapter is similar to that discussed in Chapter 2 (Section 2.3). Multinomial logit models are estimated for each of the 14 cohorts of university graduates. We consider a five-way categorical dependent variable: i) employment (WORK), ii) professional training (TRAIN), iii) higher degree study (STUDY), iv) unemployment (UN), and v) out of the labour force (OLF). Compared to Section 2.3, permanent and temporary employment have been aggregated into a single category. We decided to abandon this distinction because temporary employment represents only a minor share of the employed (less than 5% on average across the 14 years) with not enough movement over time to make its separate consideration worthwhile.[6] We are aware that considering the employed as a homogeneous group of individuals misses out important aspects concerning the 'quality' of the jobs graduates do. As stated in the introduction, this issue will be dealt with in Section 3.7.

The control variables used in this chapter are a subset of the variables used in the analysis on 1993 graduates and defined in Appendix 2A. For instance, department-level information is not used because it is unavailable for earlier cohorts. Moreover, some variables were aggregated into larger groups like university-type dummies which were replaced by the variable LEADUNI taking value 1 if the individual graduates from a restricted group of leading institutions and zero otherwise.[7] This variable is expected to be informative of potential 'cream-skimming' effects linked to the reputation of the institution attended. For related reasons, a similar level of aggregation was also adopted for degree performance (POORDEG equals 1 if student degree class is a 2.2 or below, and zero otherwise), and social class (SCLOW equals 1 if parental occupation is SC IIIM or SC IV-V, and zero otherwise).[8]

---

[6] The proportion of graduates in temporary employment (TWORK) has consistently risen over the 14-year period, but at a very slow rate.

[7] These institutions were chosen among those featuring consistently in the top spots of UK universities' official league tables. We also tried to ensure that all fields of study were represented adequately within this highly selected pool of universities. For reasons of confidentiality the institutions are unnamed.

[8] This variable is a three-way categorical variable: i) SC I+SC II, ii) SCLOW, and iii) SC IIINM+SC OTH. The aggregation of social class groups can be further justified with the evidence

3.4 Sample and summary statistics

The chapter exploits USR student-level information on complete cohorts of students who graduated in the UK in the period 1980-1993.[9] The 14-year span between the oldest and youngest cohorts enables us to have a feel for medium-to-long term changes in the determinants of first destination decisions by new university graduates. As in Chapter 2, non-respondents, medical and overseas student were excluded from the sample. In addition, graduates with entry qualifications other than A-levels or Highers were dropped owing to the much smaller number of graduates with 'non-traditional' entry qualifications in the early 1980s. Descriptive statistics on these 'non-selected' groups of students are shown in Figure 3.4. The proportion of graduates without A-level qualifications has more than doubled between 1984 and 1993. The proportion of medical graduates has been remarkably stable over the whole period. Survey non-response rates have remained in the region of 12% until 1990, followed by an isolated one-off peak in 1991, and by a steep decline in 1992 and 1993.[10] The two humps corresponding to the recession years suggest that non-response is counter-cyclical. After selection, we are able to use samples ranging from 45,600 to 56,200 graduates in any year.

---

from the analysis conducted in Chapter 2 that, with the exception of working classes, we generally found weak effects associated to most social class dummies.

[9] As already pointed out in Chapter 2, the 1993 cohort represents the most recent wave of graduates for which individual data are available.

[10] Throughout the sample period, relatively higher non-response rates tend to be found among males, low achievers, Art graduates, married individuals and mature students.

3.4.1 Changes in first destinations

Figures 3.5-3.9 show the trends in graduates' employment, professional training, higher degree study, unemployment, and non-employment both by year and gender. The patterns of employment and unemployment measured as the proportions of graduates in each of these groups are consistent with the trends in the wider labour market: two recessions in the early 1980s and early 1990s intercalated by a period of steady expansion during the mid and late 1980s. The proportion of graduates who reported to be inactive[11] six months after graduation increased steadily during the 1980s reaching a peak in 1990 followed by a steady decline in the early 1990s. The proportion of graduates engaged in professional training dwindled during the 1980s, showing signs of recovery in the 1990s. Finally, graduate numbers in higher degree courses have remained stable until the late 1980s after which they rose steadily.

It is clear from these figures that the first destination pathways of male and female graduates have evolved differently over time. In 1980 the proportion of female graduates in employment six months after graduation was 12.5% lower than for males (Figure 3.5). This gap gradually narrowed during the 1980s, especially after 1988, when male employment declined twice as fast as females'. As a result of these trends, the gap disappeared in 1991 and for both 1992 and 1993 females overtook males. Figure 3.5 shows that in the 1990s unemployment among male graduates has grown at a significantly higher rate compared to females. Overall, male graduates seem to have been affected more than females by the recession of the early 1990s. Different patterns also emerge with respect to professional

[11] This category consists mainly of individuals taking 'time out' to travel.

training destinations, where the observed decline during the 1980s is almost entirely driven by the fall in professional training undertaken by female graduates.[12] The 18% gender gap in the proportion of graduates undertaking professional training has gradually narrowed during the 1980s reaching its lowest level in 1989 and stabilising around 9% in the early 1990s. This is mainly due to the decline in the number of graduates undertaking teacher training, which has traditionally attracted more female graduates. Finally, more males stay on to study for higher degrees, whilst more females take time off after graduation. However, unlike for the employed and unemployed, the male-female gap with respect to the other first destination has remained rather stable over time. For instance, male graduates have a higher propensity (5.5% higher on average than females) to undertake higher degree studies throughout the sample period. This is partly due to the persistence of gender differences in subject choice, and particularly the high concentration of men in courses with a high 'stay-on rate' like Chemistry and Physics.

3.4.2 Changes in subject choice

Table 3.1 shows that the distribution of graduates across the range of courses considered in this study is similar for the 1980 and 1993 cohorts, suggesting little movement over time. Overall, between 1980 and 1993 the proportion of female graduates has increased from 39% to 46%. These figures are in line with OECD data on the larger population of university graduates (43% in 1982 compared to

[12] USR first destination information reveals that in 1980 57% of female trainees engaged in teaching training compared to 45.4 % in 1985 and 36% in 1990.

47% in 1992) (OECD, 1993, p. 23). Consequently, the proportion of women has increased in almost all courses, with the exception of Computer Science. More women appear to graduate in male-dominated courses like Chemistry, Physics, Business, Economics, and Engineering. On the contrary, subjects which are traditionally female-dominated like Education, Modern European Languages, and Classics have not only remained largely a preserve for women, with little change over time, but have in places become even more 'feminized' as in the case of Allied Medicine (ALMED).[13]

3.4.3 Changes in other student characteristics

Table 3.2 shows summary statistics for other key variables for each of the years from 1980 through to 1993 by gender. The proportion of graduates from the selected group of leading universities is about 10% of the population, is slightly male-dominated and has remained fairly stable over time. 'Working class' (SCLOW) graduates account for less than 20% of the graduate population, and the trend shows a decline during the 1990s.[14] The proportion of graduates educated at an independent school has nearly doubled between 1980 and 1993, while the proportion of graduates with 'good' degrees (DFIRST+DUSEC) has increased remarkably during the period 1980-1993, especially for women. The overall

[13] Brown and Corcoran (1997) report similar evidence for the US. Between 1968 and 1991 significantly more female have enrolled in 'male' majors like Business, Engineering and Physical Sciences, while typically 'female' majors like Education and Health-related degrees have not experienced comparable inflows of men.

[14] These trends may reflect to an extent the change in the wider social class structure in UK. The shift from manufacturing to service has had the effect of reducing the proportion in the population employed in skilled manual occupations. Wolf (2002, p. 49), using LFS data shows a decline from 18% to 12% for craft and related occupations, and an increase from 29% to 37% in managerial, professional and technical jobs.

growth in the proportion of good degrees was 75% and 58% for females and males, respectively. We notice, however, similar trends with respect to A-level performance. The proportion of graduates with at least 24 points overall in the sum of their 3 (5) best A-level (Higher) scores (BBB+) has increased by around 40% for both males and females. A first explanation for the trends in degree class is that the average ability level of the student intake has increased over time in response to a more stringent admission policy set by universities to cope with the growth in the number of applicants. Secondly, degree performance may have improved because students having perceived the importance of a good degree in an increasingly competitive labour market have become more motivated and dedicated in their studies. Finally, it is also possible that awarding standards have simply changed over time.[15]

3.5 Estimation results for all students

This section presents the results from the multinomial logit regressions run separately for each of the fourteen cohorts of graduates who left university between 1980 and 1993. The results refer to marginal effects calculated at the means of the explanatory variables. Given the emphasis on trends, the results are largely presented in a graphical format. Figures 3.10-3.15 trace out the impact of some key variables on first destination probabilities over time, providing also information on the 'within-cohort' and 'between-cohort' statistical significance of the estimates. In these figures, in any year, a filled (hollow) point indicates that

[15] Johnes and McNabb (2002) argue that during the 1980s there was significant grade inflation in higher education in the UK. If there was grade inflation, this had probably occurred prior to

the effect is statistically (insignificantly) different from zero.[16] Finally, to gauge the sensitivity of the estimated effects to the business cycle, Table 3.3 reports the correlation coefficient between the marginal effects of some key variables and an indicator of the cyclical component of the business cycle based on the UK real GDP in the period 1980-1993.[17] To aid interpretation, a high positive (negative) and statistically significant correlation coefficient between the business cycle and a variable's marginal effects on the predicted probability of employment (unemployment) indicate that the impact of that variable on graduates' employability is pro-cyclical.

3.5.1 Institution effects

Figure 3.10 shows the marginal effects of graduating from a leading university on the probability of employment, unemployment and further study. Interestingly, during periods of economic downturn graduates from top-ranked universities are significantly less likely to be still unemployed, even after controlling for other key factors like degree performance, subject studied and social class. Table 3.3 shows that the correlation between the business cycle and the employability *premium* of graduating from a leading institution (defined as lower probability of being unemployed) is positive and statistically significant (0.60). This confirms that the

---

university entry.

[16] Statistical significance (5% level) is based on the standard errors of the estimated coefficients from which the marginal effects are derived. We also performed log-likelihood ratio tests on individual coefficients (see Appendix 3A for details) to test whether the differences in the estimated effects between too consecutive years were different from (equal to) zero. We found that the tests widely rejected the null hypothesis of equality of the coefficients of interest between consecutive years at the 5% level.

'LEADUNI' effect is strongly counter-cyclical, in the sense that the 'cream-skimming' of students from leading institutions intensifies when the labour market is slack. Cost-effective recruitment strategies may induce employers to concentrate on a smaller pool of universities with an established reputation or a record, based either on previous experience or on prejudices, for having supplied successful graduates in the past. In the early 1980s the positive screening effect of graduating from leading universities is also substantiated in a significant employment *premium*. However, between 1987 and 1992 LEADUNI graduates are significantly and increasingly less likely to be in an occupation, especially in the early 1990s. These trends seem to be largely explained by the (nearly) symmetric increase in the probability of further study. The latter result may reflect the fact that leading institutions typically have a large postgraduate population and undergraduate students are more exposed to different aspects of postgraduate education, including a greater awareness of the programmes' content and the career prospects they offer.[18]

3.5.2 Gender effects

Figure 3.11 shows the pattern of the marginal effect of being male on the probability of employment, further study and unemployment. In the early 1980s, the hypothetical male graduate with sample mean characteristics was nearly 6

---

[17] We thank Prof. Marco Gallegati for making the data available to us. In Gallegati (2002) the series is obtained by applying an approximate band-pass filter to data from the OECD Business Sector Data Base.

[18] Leading universities are traditionally research-oriented and are able to attract a greater amount of public funds from the Research Councils. This implies, *inter alia*, better funding opportunities for postgraduate study and greater incentives to stay on.

percentage points more likely to be employed than his female counterpart. From 1981 the employability *premium* of being male has steadily declined and after disappearing in 1990 it has become negative in the period 1991-1993. In 1993, a male graduate was about 3.5 percentage points less likely to be employed than the otherwise 'observationally identical' female.

With respect to unemployment, the marginal effect of being male (relative to female) follows a pro-cyclical pattern. Table 3.3 reveals that the correlation with the cycle is negative and statistically significant (–0.79), which implies that males are more likely than females to be unemployed when demand is slack. It is striking how in the early 1990s the employment prospects of men have significantly deteriorated vis-à-vis women, even after controlling for subject studied and degree class. In the next section we analyse more closely the possible reasons for these diverging trends after 1989 by estimating separate models for males and females.

3.5.3 Degree class effects

Figure 3.12 shows the pattern of the marginal effects of graduating with a 'poor' degree class (2.2 or lower) *vis-à-vis* a 'good' degree (2.1 or higher) on the probability of employment and unemployment. Evidence from a study enquiring about career-related perceptions of university students (Connor *et al.*, 1997) suggests that upper second class honours are increasingly regarded as a threshold below which many recruiters do not shortlist candidates. It is interesting to note that the probability that low achievers experience unemployment after graduation is systematically higher during periods of economic contraction. This is confirmed by a statistically significant correlation coefficient of –0.88 between the marginal effect of POORDEG on the probability of unemployment and the business cycle (Table 3.3). This result suggests that 'cream-skimming' by academic ability is more likely to occur when the labour market is slack. Table 3.3 also indicates that low achievers become increasingly less likely to undertake postgraduate studies than high achievers in periods of adverse economic climate. We note that the gradient of the effects has become steeper in the early 1990s. This could be partly related to the fact that degree performance has improved over the years and so has the negative signal that poor performance conveys to employers.

3.5.4 Socio-economic effects

Figures 3.13 and 3.14 show the pattern of the marginal effect of coming from a working class family (skilled manual, partly skilled and unskilled parents) and of being educated at an independent school, respectively. The type of school attended has been used in related literature as a proxy for social background (Johnes and Taylor, 1990). The results suggest that working class students are generally more likely to be unemployed, especially during periods of economic downturn (the correlation with the cycle is –0.82). It is also evident that the SCLOW effect was stronger in the early 1980s. For instance, in 1982 a working class graduate was 3.2 percentage points more likely to be unemployed and 4.4 percentage points less likely to be employed than the observationally identical individual from SC I and SC II. In 1992, the unemployment gap was 1.8

percentage points, while differences in employment were no longer statistically significant. Socio-economic factors, although less influential than in the past, still positively affect graduates' employability, particularly during an economic recession.

The pattern of the independent school effects (INDEP), shown in Figure 3.14, tells a similar story. In 1982, a graduate educated at an independent school was 5.2 percentage points more likely to be employed and 4.7 percentage points less likely to be unemployed vis-à-vis the otherwise 'observationally identical' former LEA pupil. Both the magnitude and the statistical significance of these effects have steadily fallen during the 1980s, particularly with respect to the probability of employment. Interestingly, the effects have resurfaced in the early 1990s. Furthermore, Table 3.3 shows that the correlation coefficient between the business cycle and the independent school effect on the probability of unemployment is 0.62 and is statistically significant. This means that during economic downturns the 'bonus' of coming from an independent school education background in terms of facing a lower probability of unemployment is higher. These results seem to support the view that social networks, defined as social ties to individuals in high-paying jobs, can represent a cost-effective recruitment strategy and also a useful screening device. In fact, since social ties tend to occur among persons with similar attributes, adverse selection in the labour market will push employers to solicit referrals from high-ability employees (Doeringer and Piore, 1971).

3.5.5 Degree course effects

Table 3.4 shows the marginal effects of subject studied on first destination probabilities for just the two cohorts of 1980 and 1993.[19] For the majority of courses employability *premia*, both in terms of higher probabilities to enter an occupation and lower probabilities to be still unemployed six months after graduation, have narrowed vis-à-vis Humanities (default).[20] This evidence does not seem to accord with the claim that the graduate labour has become increasingly stratified along university course lines (Lynton, 1993). This tendency is particularly evident for Mathematics and Engineering, and more mitigated for Computer Science, Economics and Business graduates. Chemistry and Physics graduates have also experienced a dramatic fall in the probability of employment relative to Humanities, but the decline is almost entirely counter-balanced by the outstanding increase in the probability of postgraduate study. We also note some important exceptions. Relative employment and unemployment *premia* have changed very little in Biology, Sociology, Modern European Languages, Allied Medicine, and in the case of Creative Arts and Education, employability has increased vis-à-vis Humanities.

To gain an idea of how these effects have evolved over time, Figure 3.15 shows the dynamic patterns of the marginal effects of some degree courses on the predicted probability of unemployment. The graphical analysis suggests that graduates in vocational-oriented courses like Allied Medicine and Law become

---

19 For presentational purposes, the effects on the probability of OLF (omitted category) are not reported. However, as the sum of the marginal effects must add to zero, the omitted effects can be easily calculated.

20 Humanities has been chosen as the default category because of its relatively large cell size. Besides, male and female graduates are represented in similar proportions, i.e. there is very little gender bias in the graduate population.

increasingly less likely to experience unemployment vis-à-vis Humanities graduates as the economy slows down. In other words, the marginal effects of these courses are markedly counter-cyclical, as confirmed by their high and positive correlation with the business cycle reported in Table 3.3 (0.72 and 0.81, respectively). Counter-cyclical marginal effects are also found for Economics and Engineering, although these are significantly less pronounced especially in the early 1990s.

We note that gender is a potential discriminating factor between subjects like Mathematics, Engineering, Computer Science, Economics and Business whose graduates have seen a deterioration over time in their relative employment prospects and subjects like Biology, Sociology, Modern European Languages, Allied Medicine, Creative Arts and Education where graduate employability remained stable or even improved. In fact, not only are the courses from the first (second) group traditionally male-dominated (female-dominated), but gender concentrations in these subjects have also undergone little change over time compared to other courses (see Table 3.1). Alternative identification criteria like, for instance, scientific *versus* non-scientific or vocational *versus* non-vocational degrees are less clear-cut. In fact, Biology and Allied Medicine are undoubtedly scientific, while Allied Medicine, Creative Arts, and Education, and to some extent Modern European Languages, are perceived as vocational courses. Therefore, the different labour market performance of male and female graduates especially in the early 1990s may be related to gender differences in subject mix. Because we do not control for graduates' type of employer and/or occupation, subject effects are expected to encapsulate shocks specific to the sector and/or the occupation that graduates go into. This is particularly true for courses geared to specific occupations or sectors. For instance, the relative deterioration of the employment prospects of Engineering graduates is likely to be largely due to the heavy crisis that hit Engineering and Construction in 1990-1992, causing a 10% contraction of the sector.[21] In the next two sections we will investigate in more depth the link between gender, subject studied, sector of employment and type of occupation to shed additional light on employability *premia* by subject studied.

3.6 A gender-specific analysis

The significant differences in the early labour market performance of male and female graduates illustrated in Figures 3.5 and 3.6, as well as the 'diverging' trends in the MALE effect in the 1990s discussed in Figure 3.11, led us to re-estimate the model for men and women separately. An interesting question is to assess to what extent these trends are the consequence of changes in observable characteristics like subject mix and degree performance, rather than the consequence of unobserved factors such discrimination and/or occupational and sectoral segregation. As Brennan *et al.* (1993) note, '[There is an ongoing interaction between genders, motivations, career orientations and choice of study, the effect of which carry over into the labour market. These are long-term, well established cultural patterns, which are not easy to shift.]' (p. 21).

In Chapter 2 (Section 2.7) we found that in 1993 the higher unemployment rate among new male graduates relative to their female counterpart was almost entirely

---

[21] D. Grow, 'Recession in Engineering worse than 1990' The Guardian, Wednesday, 3 October 2001.

due to unobservable factors. One tentative conclusion drawn from that result was that the unexplained gap could be partly due to affirmative-action recruitment practices favouring females. In this section we aim to extend to previous years the decomposition of gender differences in first destination probabilities already used in Chapter 2. Table 3.5 shows the results of the decomposition of the gender gap in the probabilities of employment and unemployment by year.[22] For instance, when male coefficients are used as standard, 63.3% of the 12.6 percentage point (positive) employment gap between male and female graduates observed in 1980 is explained by gender differences in factors directly accounted for in the data. Likewise, the 5.8 percentage point unemployment gap observed in 1993 remains almost entirely unexplained by the model. Between 1980 and 1990 gender differences in observable characteristics explain a substantial proportion of the gap, especially with respect to the employment *premium* for men. This evidence can be explained by the fact that over time more women attain their degree in highly 'marketable' courses traditionally dominated by men such as Business, Economics and Engineering, particularly in those years characterised by the expansion of the Banking, Finance & Insurance sector, on one hand, and by a shortfall of engineers, on the other.[23] Furthermore women's degree performance measured as the proportion of 'good' degrees has improved faster than men's especially after 1988 (see Table 3.2). This may have contributed to enhance women's employability vis-à-vis men.

---

[22] We also tried to decompose inter-cohort differences in predicted probabilities into the portion attributable to differences in observable cohort-specific characteristics and the portion due to unobservable factors. We found that differences in 'unobservables' (different labour market conditions) explain, on average, over 90% of the total gap, even when characteristics were 'swapped' across cohorts graduating in similar or comparable phases of the business cycle.

Table 3.5 also shows a dramatic drop in the explained percentage of the gender gap in both employment and unemployment from 1991. This 'break' in the series observed for both WORK and UN is only partly reflected in the fall of the 'goodness of fit' in the gender-specific regressions (pseudo-R^2), and it is even more striking when contrasted with the absence of breaks in the corresponding series for the other first destination categories (Table 3.6).

There is no obvious single explanation for the trends discussed in Table 3.5. Rather, these results are likely to arise from a combination of several factors. A first possibility is that affirmative-action recruitment policies favouring females may have become more widespread in the 1990s.

A second explanation could be the tendency, common to most courses, towards a lesser degree of 'labour market specialisation', coupled with the persistence of significant gender segregation by type of employer and/or type of occupation through the 1990s. The argument goes as follows. If, on one hand, more graduates find employment in sectors and/or occupations not strictly related to the content of their courses, subject-specific effects become a weaker proxy of the impact that sector-specific shocks have on graduates' employability. If, concomitantly, gender segregation remains unscathed in some sectors and/or occupations, the dramatic fall in the explained portions that emerge in Table 3.5 is caused by unobserved sector-specific effects, which remain increasingly unaccounted for by differences in the subject mix. As a crude test for this line of argument, Table 3.7 shows, for some courses,[24] the modal sector of employment and the modal occupation with

---

23 This reflects the general upsurge in demand in construction and engineering, but also the short-term effect of the Channel Tunnel project.

24 Here we consider only those courses that are generally perceived as being geared to a narrowly defined range of labour market segments.

the corresponding proportions of graduates employed in these sectors/occupations. The proportion of Engineering graduates who find employment in the Engineering & Construction sector (EC) has declined by 12 percentage points between 1988 and 1992, but the fall was twice as fast between 1990 and 1992. These figures reflect the crisis that badly hit the Engineering sector between 1990 and 1992. Similar trends apply to Economics where the proportion of graduates going into Banking, Finance, & Insurance (BFI) has fallen by 10 percentage points between 1988 and 1992, four fifths of which between 1990 and 1992 only. With the exception of Education and Allied Medicine degrees, both the degree of sectoral and occupational specialisation has declined for all courses, especially after 1990.[25] On the other hand, Table 3.8 reports the differences in the distribution of male and female graduates across the main employment sectors/occupations. Gender segregation is still significant in sectors like EC, Public Administration, Health & Education (PHE), and Commercial & Allied Services (COMM), although there are signs of decline over time. In the light of this joint evidence, we conclude that the weaker link between subject studied and sector/occupation of destination in the 1990s is probably one of the reasons why gender differences in early labour market performance remain largely unaccounted for by the model.

Finally, an additional explanation could be related to unobserved gender differences in the 'quality' of employment. For instance, females may be less reluctant than males to accept jobs for which they are overqualified. This higher

[25] An interesting case is represented by Physics courses. As a result of the crisis of engineering, EC is no longer the modal sector of employment for Physics graduates in 1992. However, despite the modal sector has changed, the degree of specialisation, measured as the proportion of graduates employed in the modal sector, has remained fairly constant over the 4-year period.

'flexibility' may facilitate the transition from university to work in periods of economic downturn when there are fewer jobs available and the competition for 'graduate' level jobs, crudely defined as occupations for which a degree is required, becomes fiercer. In a country like the UK where graduate numbers have risen steadily during the last two decades, bottleneck effects can be particularly strong. The mismatch between qualification held and job requirements is commonly referred to as 'overeducation' or under-employment and has recently received considerable attention in the literature. Although a detailed analysis of the causes of overeducation is beyond the scope of this chapter, the next section presents an empirical analysis on these issues.

3.7 'Graduate' and 'non-graduate' employment

Up to this point, no attempt has been made to account for differences in the 'quality' of the job graduates do. Arguably, when the main focus of the analysis is on time trends, this assumption is even more restrictive. Generic employment rates are no longer a satisfactory measure of labour market success because graduate employability has become more complex than the black and white picture of employment versus unemployment. Evidence suggests that graduates are taking longer to settle in the labour market with more entering jobs not specifically designated for graduates, often displacing less qualified candidates or moving into temporary employment (Pearson *et al.*, 2000).

There is also concern that the rising levels of student debt are leading students to take any job in order to be able to pay off their debt. Until the end of the 1980s, student tuition was publicly funded out of general taxation, while students' living costs were financed by a mixture of means-tested grants and parental contributions. In 1988, the introduction of the student loans, and the abolition of student maintenance grants may have exacerbated these trends.

In this section we introduce a distinction between occupations which are typically thought of as 'graduate occupations' and those 'non-graduate' jobs for which a degree level qualification is typically not required. The distinction is expected to be informative of the changes in the *quality* of graduates' starting jobs over time. Furthermore, we expect to gain additional insights into the pattern of the gender differences during the 1990s discussed in the previous section.

3.7.1 The definition of 'graduate' and 'non-graduate' occupations

A *graduate/non-graduate* job is defined according to the average number of years of post-compulsory schooling held by employees in each of the 371 occupation groups of the SOC90.[26] This information is then merged into our dataset by matching the SOC definition of jobs with the USR classification. Due to the nature of this study, we use a rather inclusive definition of *graduate* occupations. This category not only includes *traditional-graduate* jobs like professional and high-level managerial occupations which typically require 5 years of additional education after the age of compulsory schooling, but also *graduate-track* occupations like low-level management and technician jobs which require typically 3 years of post-compulsory education but not necessarily a degree. This was done because *graduate-track* occupations can be considered entry route jobs and new areas of graduate work as well as areas of work in which individuals with A-level qualifications may find employment. It is likely that over the fourteen years considered in the analysis, technology has altered the skills and qualifications requirements to perform many of these occupations. The inclusive definition of graduate occupations implies that the definition of *non-graduate* occupations is restricted to jobs like low-level clerical and manual occupations which typically require 1.5 years of additional schooling and for which a degree is clearly not necessary. Another reason for choosing a restrictive definition of non-graduate occupations is the change in the occupational coding that took place in 1990. Prior to 1990 the coding scheme was the Key List of Occupations for Statistical Purposes (KOS), while from 1990 onwards occupations were coded according to the SOC. The method used by McKnight is based on the SOC90. We are able to extend the definition of non-graduate occupations based on the SOC90 to earlier years because the USR classification has remained unchanged throughout the whole sample period. However, the match between KOS and SOC is generally fairly imprecise (Bell and Elias, 2000). Therefore, a broader definition of 'non-graduate' occupations may give rise to coding errors and possible breaks across years. Admittedly, this is an imperfect measure of quality. However, a consensus on the exact definition of what is a graduate/non-graduate occupation is hard to find and the issue still remains a main area of contention (Sloane *et al.*, 1999).

---

26 The definition of graduate occupations used in the chapter is based on a method devised by McKnight (1999), which converts qualifications held by respondents to the Labour Force Surveys

3.7.2 Descriptive statistics

Figure 3.16 shows the proportion of graduates employed in non-graduate level occupations by year and gender. During the 1980s, the proportion of women employed in non-graduate occupations was significantly higher than for men. The gender gap reached its highest point during the recession of 1981-1982, when the proportion of new graduates overqualified grew significantly, but only among females. During the economic expansion of the 1980s the proportion of overqualified females has steadily declined, while no apparent link with the business cycle is found for men. Finally, during the recession of the early 1990s the proportion of graduates in non-graduate level occupations has increased steeply for both genders reaching 22.2% and 17.5% in 1993 for women and men, respectively.[27]

3.7.3 Results

The econometric analysis is carried out by estimating logit models of the probability of entering non-graduate level, rather than graduate level (default), occupations. The analysis is conditional on employment because individuals employed in non-graduate occupations tend to have similar characteristics to the unemployed, and this would violate the IIA assumption underlying the multinomial logit approach. Furthermore, restricting the analysis to the employed enables us to control for the sector of employment. This may be important if, for instance, the overqualified are more upwardly mobile, accepting lower level jobs

into years of post-compulsory schooling.

in order to achieve a promotion or higher earnings in the future.[28] Given the less competitive environment and market rigidities of the public sector due to a stronger unionisation and more restrictive working practises, the penalty for being overqualified may be higher and, as a result, more overqualified individuals are more likely to be found in the private than in the public sector. We consider three broad sectors: public (PUB, default), manufacturing (MANU), and services (SERV).[29] Table 3.9 reports three alternative sets of marginal effects on the probability of entering a non-graduate occupation for the 1983, 1988, and 1993 cohorts. In model (a) we use the same control variables included in the multinomial logit analysis. Model (b) adds sector dummies to the previous specification. Finally, model (c) is a probit estimated by Instrumental Variables (IV) to account for the possible endogeneity of the graduate's sector of employment.[30] The employment sector of parents is used as the instrument. Previous studies for the UK suggest that 10% of young graduates were in the same occupations as their fathers and 29% in same occupational group (Chevalier, 2002).[31] The main results can be summarised below:

---

[27] The correlation coefficient between these proportions and the cyclical component of the business cycle was 0.71 and 0.46 for females and males, respectively.

[28] Occupational mobility theory (Rosen, 1972; Sicherman and Galor, 1990) predicts that overeducation represents a temporary mismatch because the overqualified are more able to move to higher level jobs.

[29] Nearly 6% of all graduates in employment did not provide information on the type of employer. Rather than dropping these individuals we decided to keep them in the sample and included in the regressions an additional dummy for the 'unknowns'.

[30] We use the *divprob* Stata command implementing Amemiya Generalized Least Squares (AGLS) estimators for probit with endogenous regressors treated as linear functions of the instruments and the other exogenous variables.

[31] There may be several reasons explaining why youngsters tend to follow into their parents' career footsteps: a) lower set-up costs, especially in occupations like farming where these costs are typically high; b) parental networking or even 'nepotism', in the sense that parents use their position to facilitate their children' career or to obtain advantages for them; c) finally, parents may transmit their ability to their offspring either through genetic mechanisms or through a transfer of their human capital.

i) Logit estimates from model (b) presented in the first two rows of Table 3.9 suggest that graduates employed in the service sector are significantly more likely to be overqualified than the otherwise 'observationally equivalent' individuals employed in the public sector. A positive but smaller effect is also found for those individuals employed in the manufacturing sector. However, when we consider IV estimates (model (c)), the effects of manufacturing and service change sign and become stronger, although they are often statistically insignificant. It is difficult to say to what extent this is due to the choice of a weak instrument (the correlation between parental and student sector of employment is on average 10%) or to an endogeneity bias;

ii) The sizeable and negative effects associated to ENGIN, MATHS, COMP, and ECON, especially in 1993, seem to provide an explanation for the declining employability *premia* vis-à-vis Humanities discussed in the previous sections. Graduates in these courses are significantly less likely to enter occupations for which they are overqualified due either to higher career expectations or to the lack of transferable skills which makes them less flexible or occupationally mobile. It is interesting to note that subject effects generally survive the inclusion of sector dummies on one hand, and are robust to alternative modelling strategies, on the other;

iii) Graduating from a leading institution significantly reduces the likelihood of being overqualified, but only for the 1983 and 1993 cohorts. Institution effects are much weaker and imprecisely estimated in 1988. These effects are robust across alternative specifications. These trends mimic the pattern of the unemployment marginal effects shown in Figure 3.10. It is likely that the same institution-specific factors that reduce the likelihood of unemployment are behind the lower probability to enter non-graduate occupations. Furthermore, these results provide additional insights into the pattern of the LEADUNI effect on employment and further study in the early 1990s. When finding a job becomes more difficult because of adverse economic conditions, graduates from leading institutions may have a relatively higher 'reservation job offer' which makes them less likely to enter 'any' occupation and more likely to stay on in higher education;

iv) With respect to gender, we find that males are generally less likely to enter non-graduate occupations. Significant gender differences are found in 1983 and 1993, but not in 1988. This may suggest that these differences follow a cyclical pattern. We also note that the gender effect survives the inclusion of sector dummies, but only when endogeneity is not accounted for. Therefore, whether or not unobserved gender differences in the 'quality' of employment may have contributed to explain the diverging employment patterns of males and females graduates in the 1990s discussed in Section 3.6 remains unclear. This result, in turn, reinforces the possibility that affirmative-action hiring policies favouring females may have been behind the observed trends;

v) Not surprisingly, graduating with a low degree class significantly increases the probability of entering a non-graduate level occupation, especially in 1993. As we already found in Section 3.5 with respect to the probability of unemployment, the gradient of degree class effects becomes much steeper in the 1990s. Furthermore, the impact of degree performance survives the inclusion of sector dummies;

vi) Working class graduates are more likely to be overqualified relative to graduates from professional, managerial and technical backgrounds, especially in 1993. Similarly, the (negative) independent school effect (INDEP) on the probability of entering a non-graduate occupation is significant and robust to the different specifications. Furthermore, the effects are stronger during economic downturns, especially during the recession of the early 1990s. It is possible that parental and social networks not only affect employability in general, but, through role model effects, also facilitate entry to graduate level jobs;

vii) Mature students are significantly less likely to be overqualified, but only in 1993. Relative to younger graduates, mature students may have higher (and self-fulfilling) career expectations and/or higher endowments of human capital either in the form of work experience directly relevant for the job or in terms of transferable skills developed earlier in their career.

3.8 Summary and conclusions

The chapter investigates the early career pathways of fourteen successive cohorts of UK graduates who left university between 1980 and 1993. We estimate cohort-specific multinomial logit equations to ascertain how the effects of factors like gender, subject studied, university attended, and degree class on graduates' early careers have changed over time, and how these changes relate to the business cycle. A number of interesting results have emerged from the analysis. First, notwithstanding the persistence of a significant stratification by subject studied in the graduate labour market, relative employability *premia* seem to have generally fallen over time. It appears that these trends are related to differences in the 'quality' of employment across subjects, i.e. Engineering and Computer Science graduate who are traditionally less likely to be overqualified for the job are those who paid the highest price in terms of falling (rising) employment (unemployment) rates in the 1990s. Second, women's employability has been less affected than men's during the economic downturn of the early 1990s. The evidence discussed in this chapter suggests that these trends are likely to originate from a combination of concurring factors including sectoral and occupational segregation, and possibly discrimination in the form of affirmative–action recruitment policies favouring females.

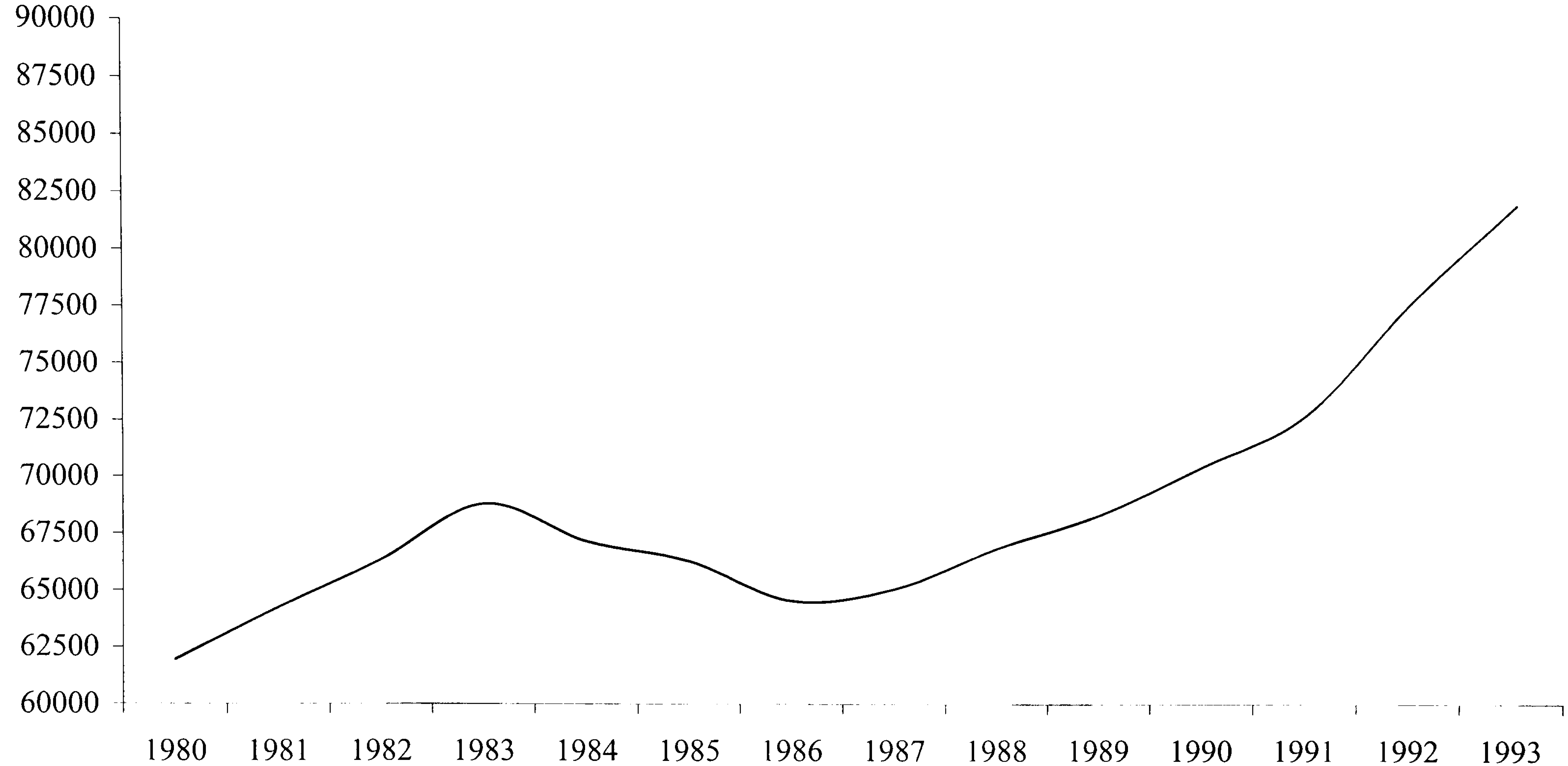
Third, graduating from top-ranked institutions reduces the probability of early unemployment, particularly when the labour market is slack, even after controlling for other key factors like degree performance, subject studied and social class. Cost-effective recruitment strategies may induce employers to concentrate on a smaller pool of universities with an established reputation or a record, based either on previous experience or on prejudices, for having supplied successful graduates in the past. Graduates from leading institutions are also more likely to undertake further study, and less likely to enter non-graduate occupations, especially during economic contractions.

Fourth, the negative consequences on graduates' employability of getting a poor degree are systematically higher when the economy struggles. However, the gradient has become much steeper during the 1990s. This is clearly related to the fact that degree performance has improved over the years and, consequently, the stratification by degree class in the labour market is bound to increase.

Finally, working class students are generally more likely to be unemployed or overqualified relative to graduates from professional, managerial and technical backgrounds, especially during periods of economic downturn, while the opposite is true for graduates who went to independent schools. This joint evidence on socio-economic factors may support the view that social networks, defined as social ties to individuals in high-paying jobs, represent a cost-effective recruitment strategy and also a useful screening device. In fact, since social ties tend to occur among persons with similar attributes, adverse selection in the labour market may push employers to solicit referrals from high-ability employees.

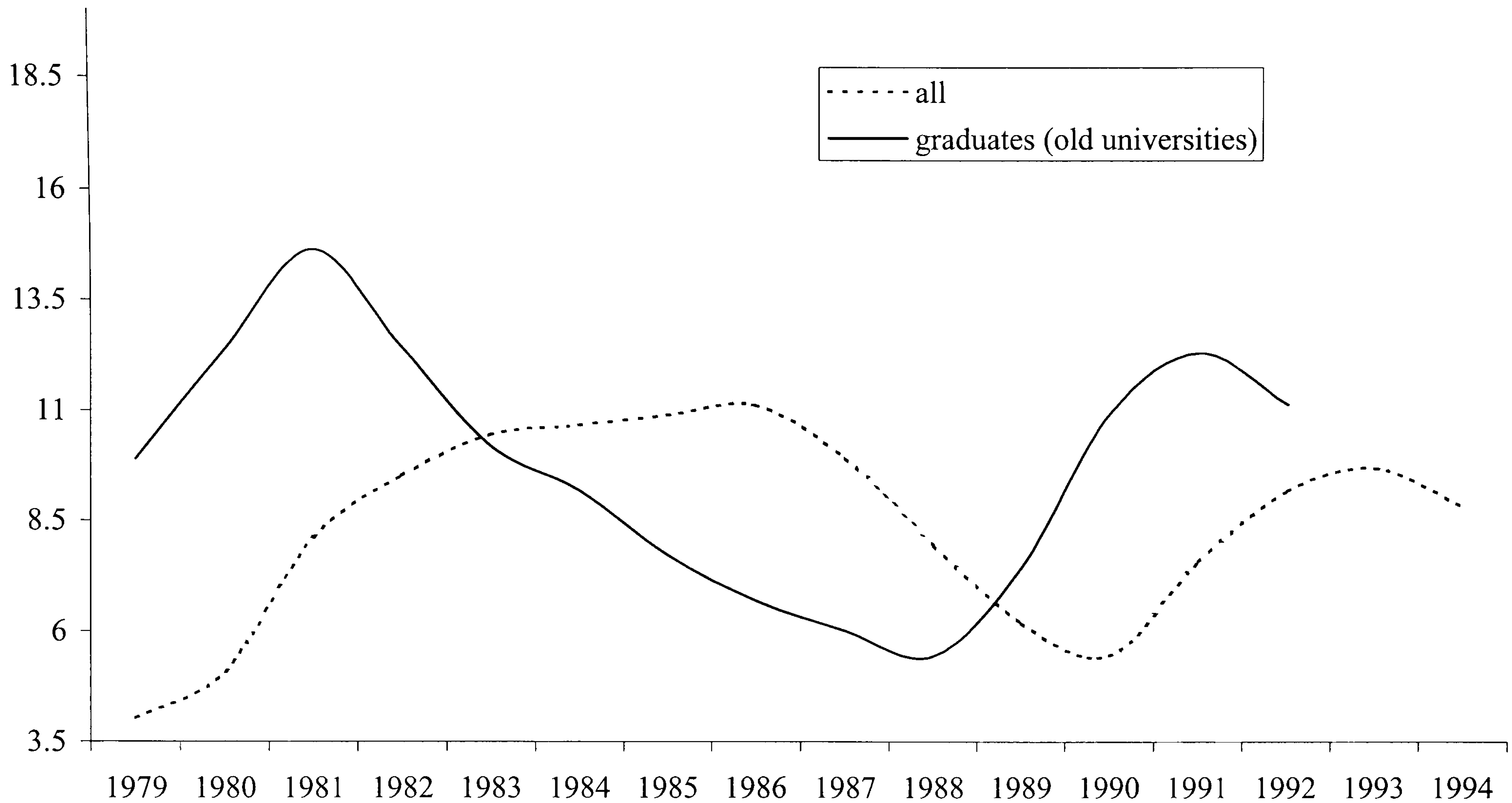
Tables and Figures

Figure 3.1 Number of graduates from 'traditional' universities



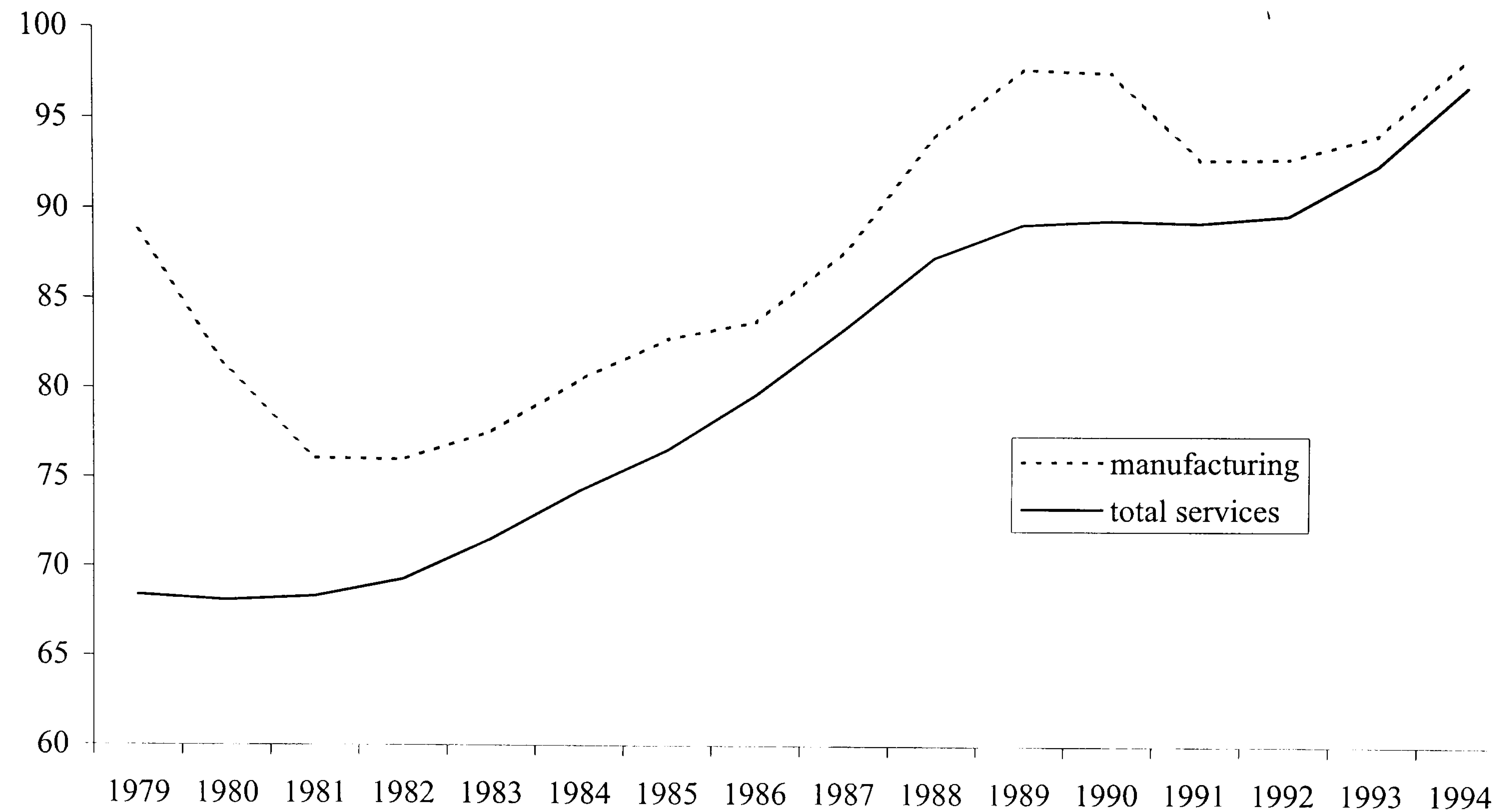
Source: USR

Figure 3.2 Trends in unemployment rates: whole workforce *vs* newly qualified graduates (%)



Source: LFS and USR (First Destination Record)

Figure 3.3 British real GDP by sector at 1995 prices (million £) (seasonally adjusted)



Source: OECD, Main Economic Indicators

Figure 3.4 Proportion of graduates in 'non-selected' groups.

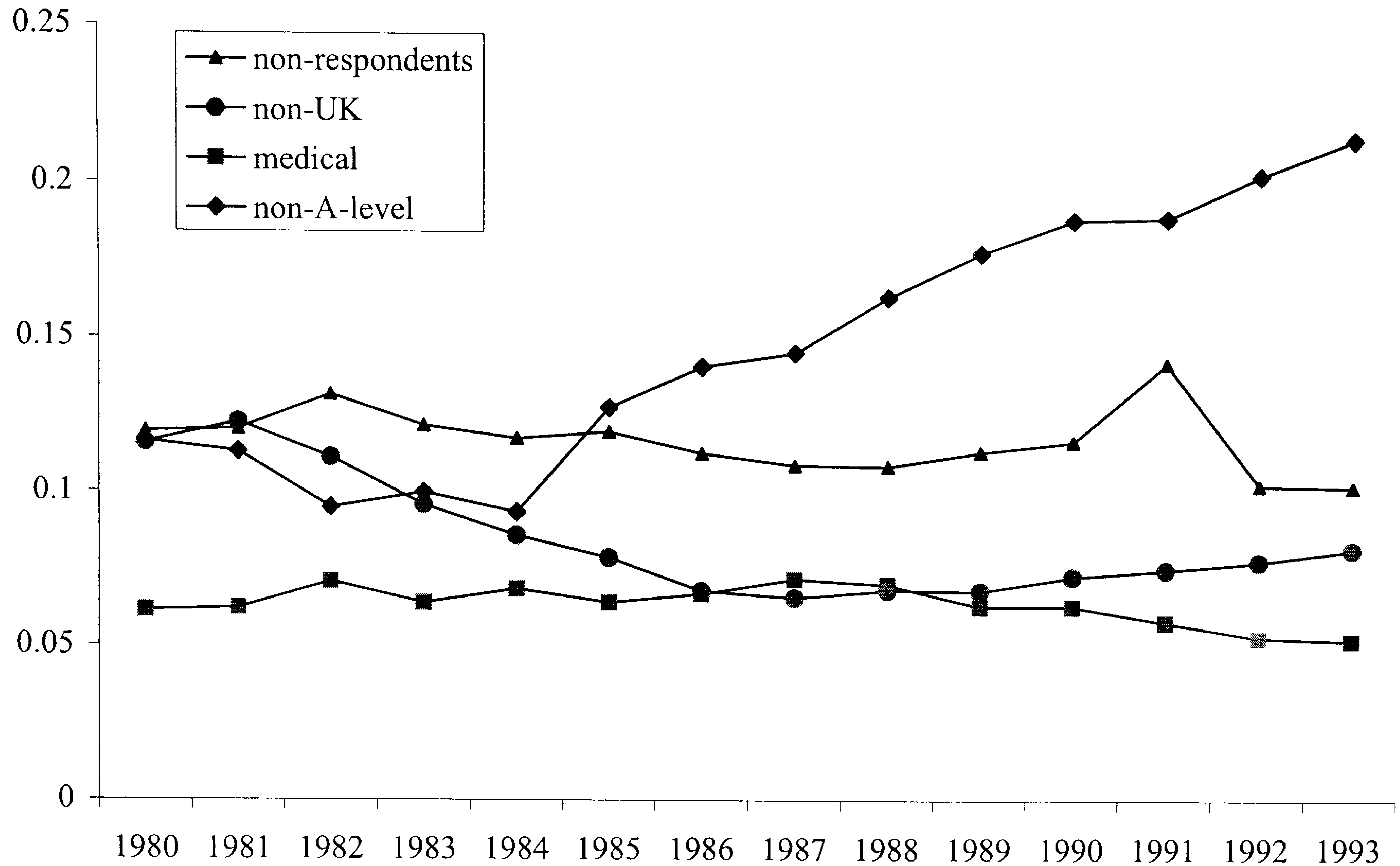


Figure 3.5 Proportion of graduates in employment (WORK) (%)

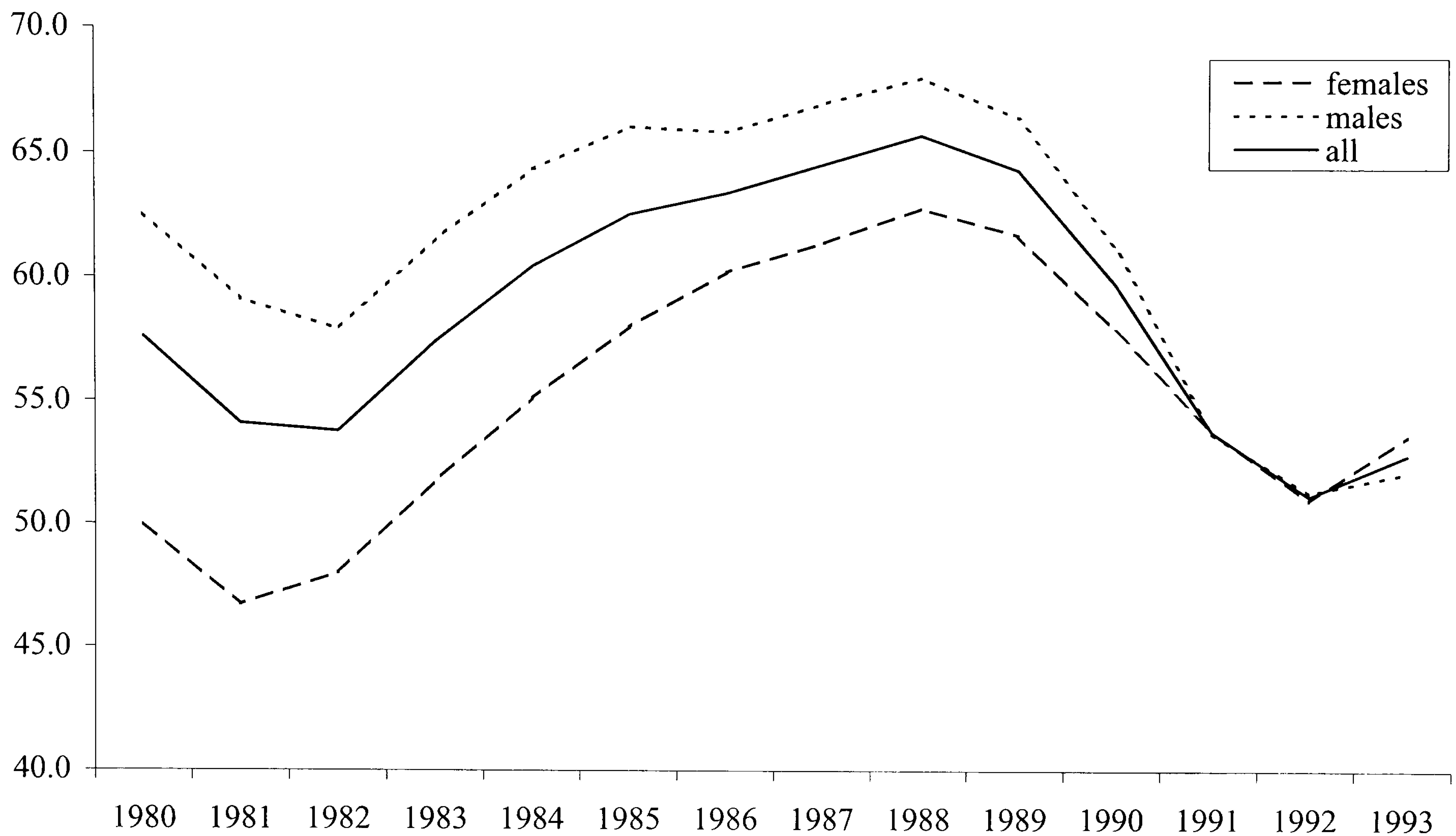


Figure 3.6 Proportion of unemployed graduates (UN) (%)

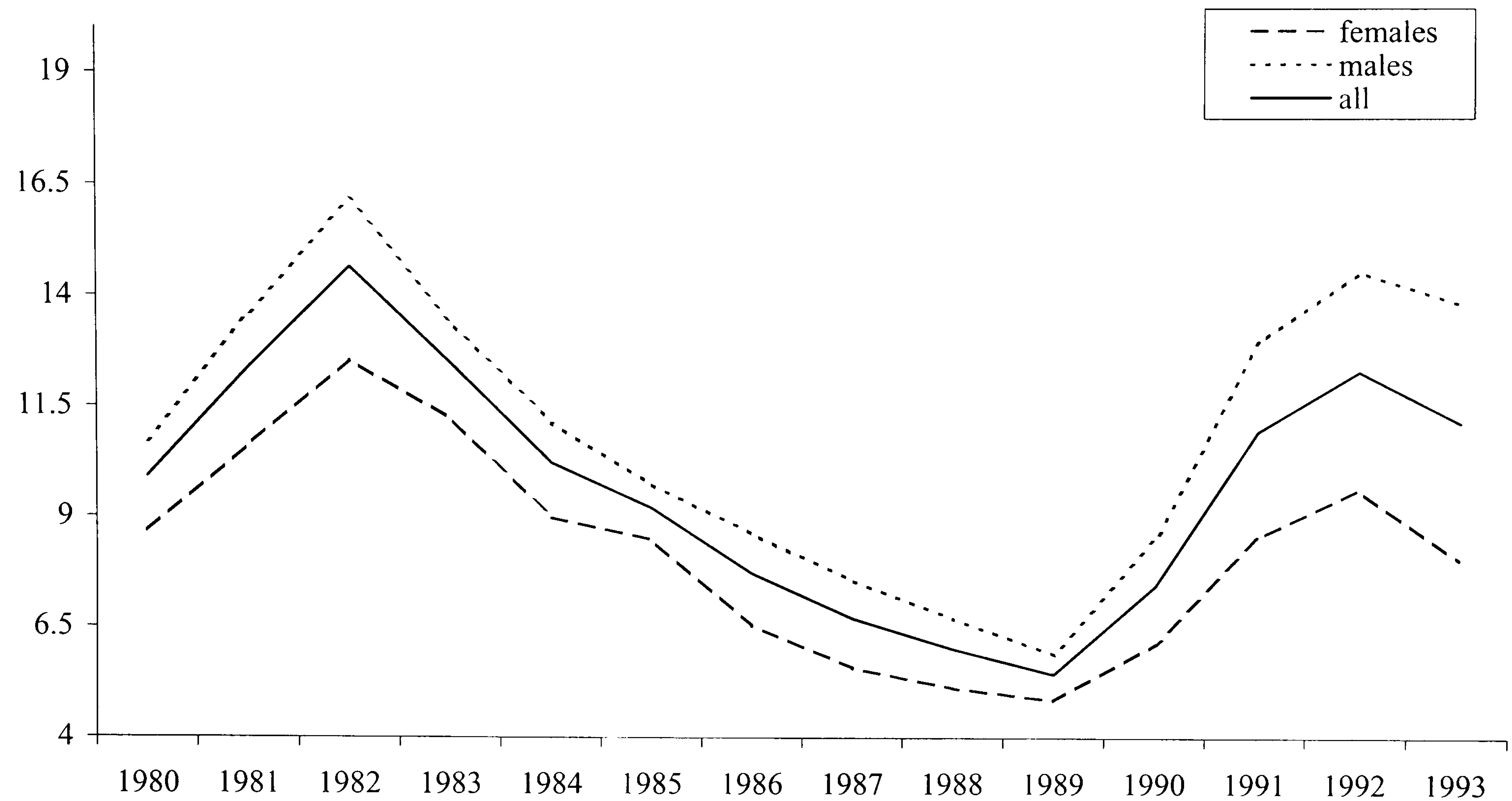


Figure 3.7 Proportion of graduates enrolled in postgraduate courses (STUDY) (%)

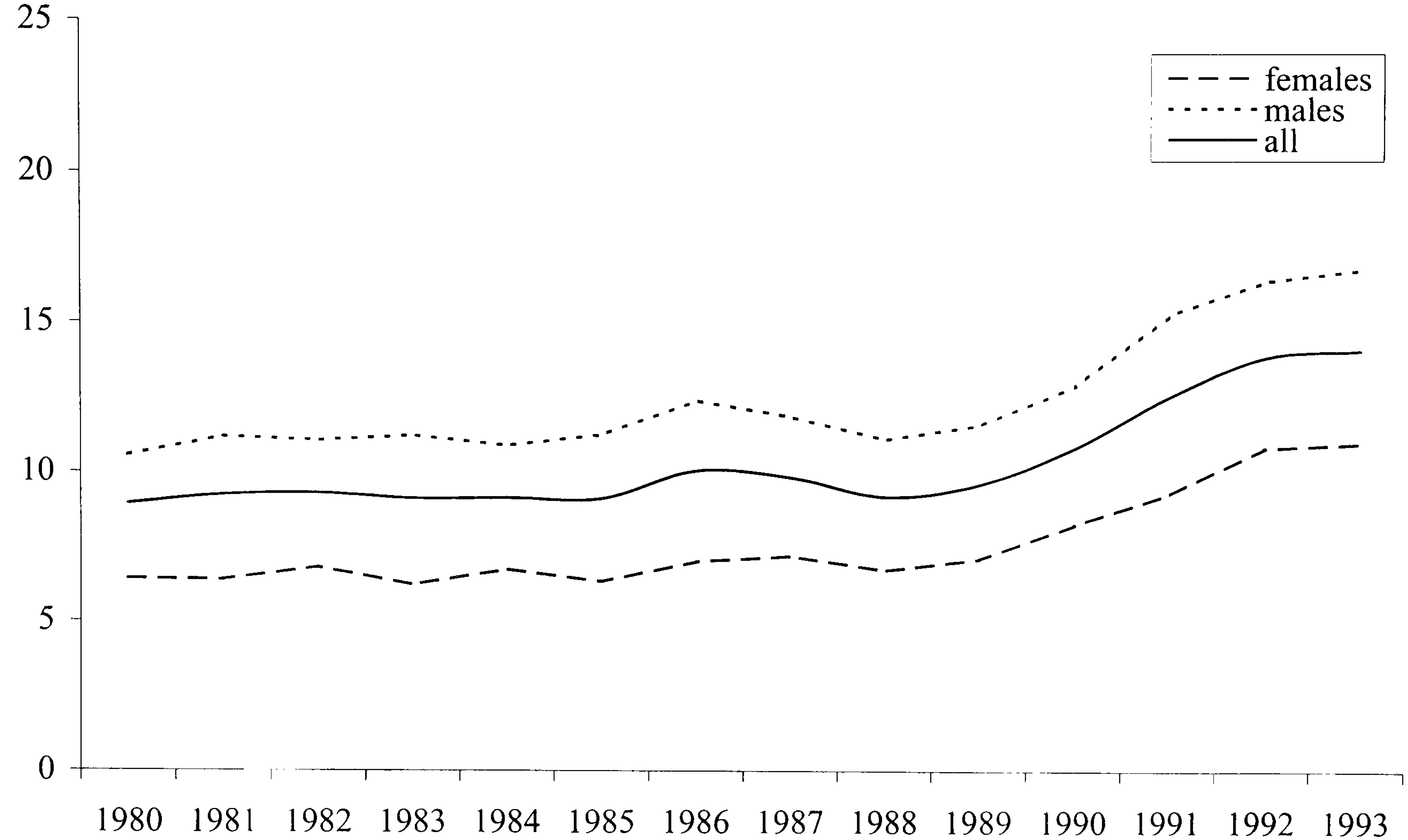


Figure 3.8 Proportion of graduates undertaking professional training (TRAIN) (%)

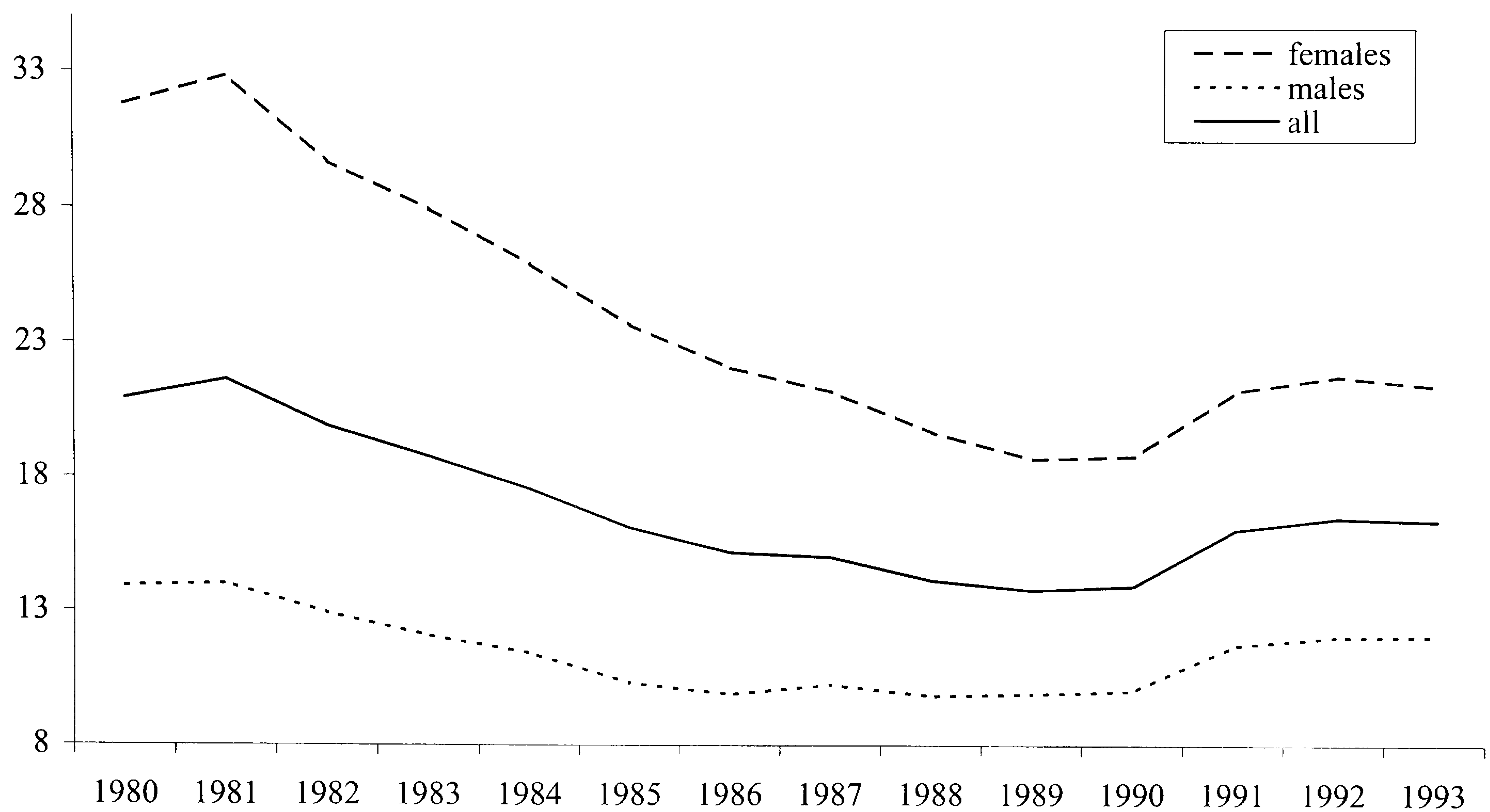


Figure 3.9 Proportion of graduates out of the labour force (OLF) (%)

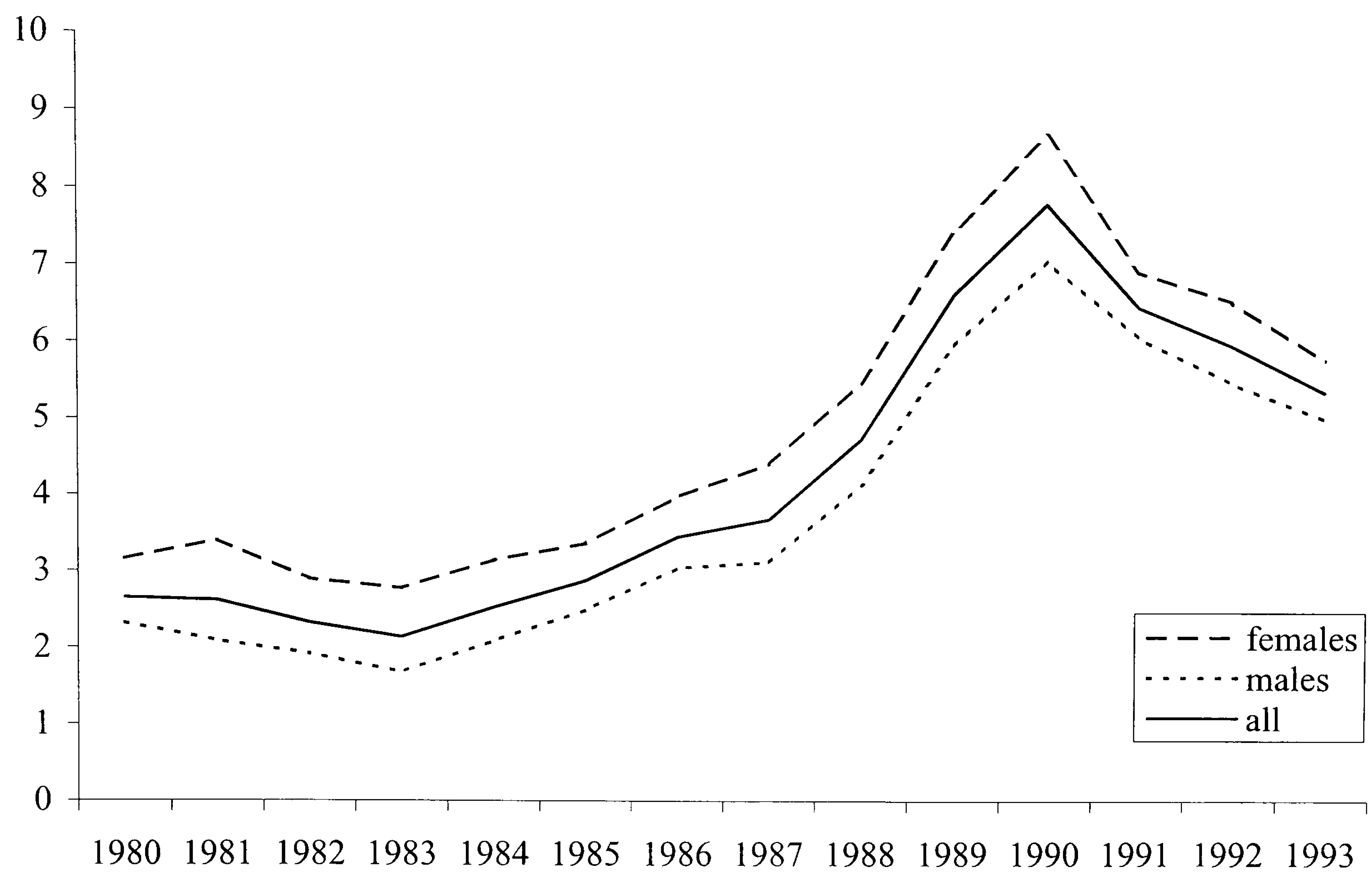


Table 3.1 Summary statistics: subject studied

	1980		1993	
	all	females	all	females
Course	%	% (within cell)	%	% (within cell)
BIOL	5.8	44.4	5.3	55.4
CHEM	3.8	22.5	4.0	35.4
PHYS	3.7	13.0	3.5	17.6
MATHS	4.5	31.0	5.6	36.8
COMP	1.3	17.8	3.0	10.9
ECON	4.7	23.0	5.6	30.2
SOCIO	2.5	60.7	2.7	69.7
POL	1.8	32.1	3.1	39.1
LAW	6.1	40.1	5.2	51.8
CLAS	7.9	65.0	7.8	67.9
ARTS	1.7	56.9	1.7	65.0
MEUL	6.4	71.2	5.3	75.1
ALMED	3.1	58.9	3.0	69.1
ENGIN	9.9	5.3	9.7	14.2
BUS	4.5	29.1	5.5	44.5
HUM	9.0	47.4	8.3	48.8
EDU	1.2	71.0	1.3	75.6
all courses		39.2		46.3
N	45584	17891	56167	26033

Table 3.2 Summary statistics: other variables (%)

	LEADUNI			DFIRST+DUSEC			BBB+			SCLOW			INDEP		
year	F	M	all	F	M	all	F	M	all	F	M	all	F	M	all
1980	6.8	14.1	11.3	36.8	36.6	36.7	33.8	30.4	32.5	16.7	20.0	18.7	10.3	15.7	13.6
1981	6.9	13.7	10.9	37.5	37.3	37.3	35.1	30.4	33.2	17.0	19.1	18.2	10.0	15.2	13.1
1982	7.3	12.5	10.3	39.3	38.8	39.0	34.7	29.8	32.7	17.1	19.3	18.4	10.5	15.3	13.3
1983	7.3	12.4	10.3	40.2	40.0	40.1	35.6	29.2	33.0	18.7	21.7	20.5	10.9	16.1	13.9
1984	7.8	12.6	10.6	43.1	42.5	42.7	38.1	31.5	35.3	16.0	17.2	16.7	10.9	16.2	14.0
1985	8.1	12.0	10.3	44.5	44.3	44.3	41.0	34.4	38.2	15.1	16.9	16.1	16.6	22.1	19.7
1986	8.5	12.4	10.7	49.2	49.1	49.2	44.8	37.6	41.7	16.0	17.5	16.8	14.5	18.4	16.7
1987	8.7	11.9	10.5	51.5	48.5	49.8	46.9	40.0	43.9	16.3	18.4	17.5	18.5	23.4	21.2
1988	8.9	11.7	10.5	53.6	51.1	52.2	47.1	39.4	43.7	16.9	18.1	17.6	20.7	25.3	23.3
1989	8.5	11.4	10.1	56.4	52.4	54.1	46.5	39.9	43.6	16.7	18.5	17.7	21.7	25.6	23.9
1990	8.6	11.8	10.4	57.2	54.7	55.8	46.8	39.6	43.6	17.2	18.8	18.1	22.5	26.6	24.8
1991	9.1	11.7	10.5	59.6	54.3	56.7	46.7	40.2	43.7	16.4	18.1	17.4	23.6	26.6	25.2
1992	8.4	11.1	9.9	61.8	55.8	58.5	47.0	40.6	44.0	15.7	17.3	16.5	23.5	26.4	25.0
1993	8.0	10.8	9.5	64.6	57.2	60.6	47.5	43.3	45.5	15.7	17.0	16.4	24.4	27.1	25.8

Table 3.3 Correlation coefficients between marginal effects and business cycle

variable	WORK	UN	STUDY
LEADUNI	-0.431	0.604*	-0.113
MALE	-0.080	-0.789*	-0.102
POORDEG	-0.330	-0.876*	0.710*
SCLOW	0.441	-0.820*	0.048
INDEP	-0.447	0.620*	0.060
BIOL	-0.109	-0.011	-0.135
CHEM	0.178	0.699*	-0.366
PHYS	0.071	0.386	-0.343
MATHS	0.145	0.391	-0.408
COMP	-0.362	0.252	0.146
ECON	-0.460	0.396	-0.332
SOCIO	-0.748*	0.293	0.405
POL	-0.638*	0.170	-0.097
LAW	-0.439	0.718*	0.463
CLAS	0.125	-0.049	0.243
ARTS	0.093	0.564*	-0.060
MEUL	0.422	0.663*	0.281
ALMED	-0.693	0.805*	-0.310
ENGIN	-0.255	0.299	-0.298
BUS	-0.596*	0.436	0.419
EDU	-0.608*	0.607*	-0.138

* denotes significance at the 5% level

Figure 3.10 Leading universities' effects (LEADUNI)

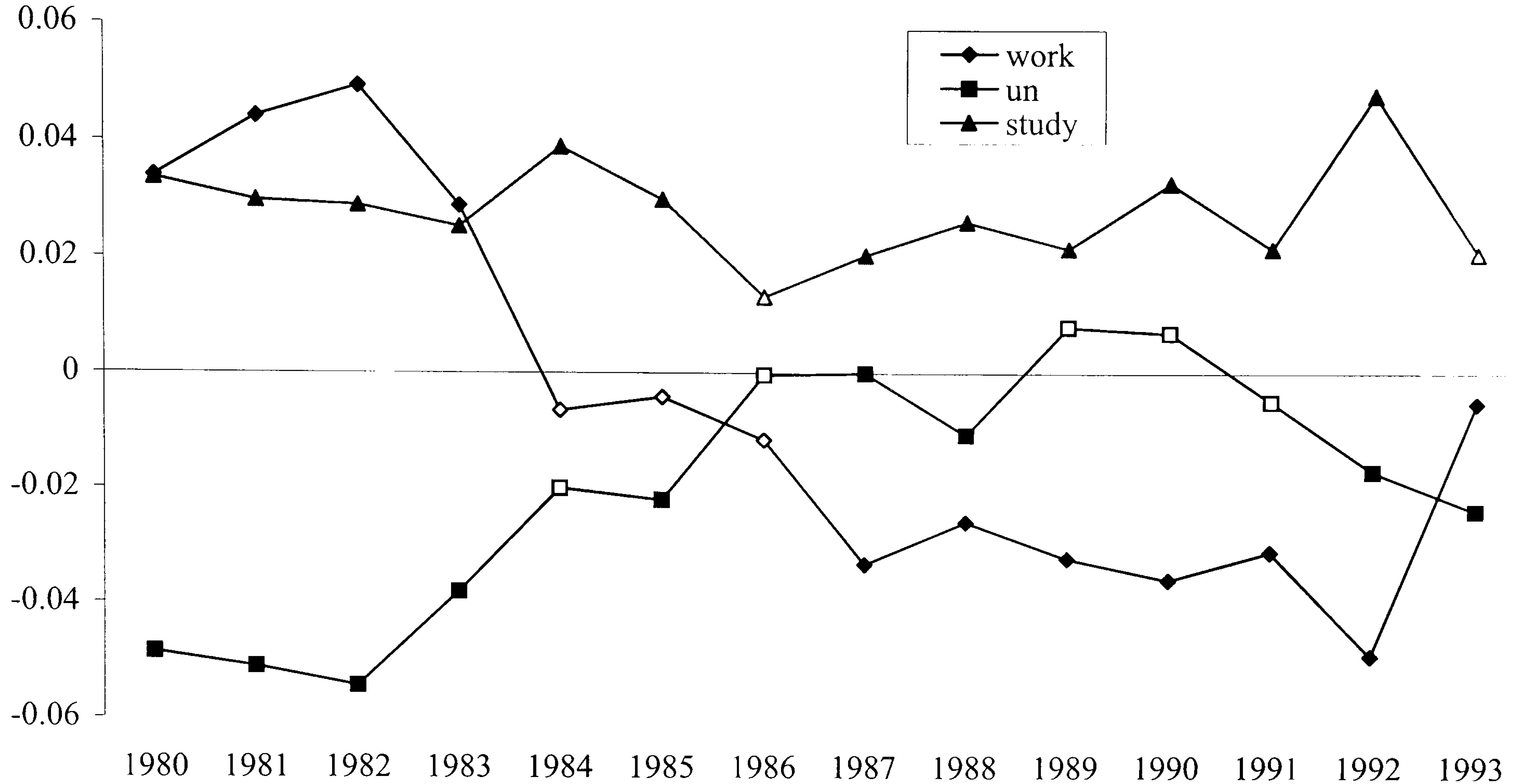


Figure 3.11 Gender effects (MALE)

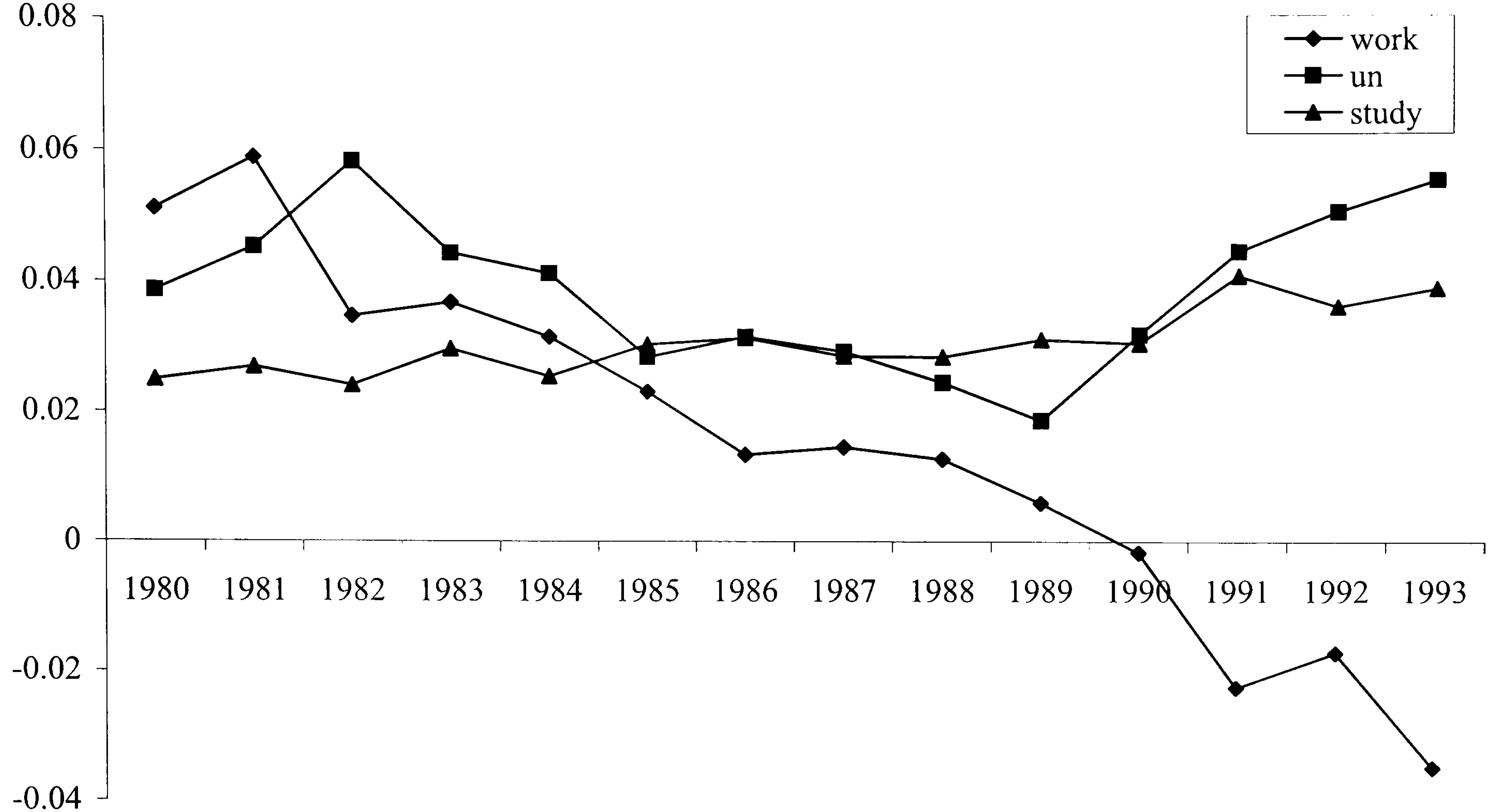


Figure 3.12 Degree class effects (POORDEG)

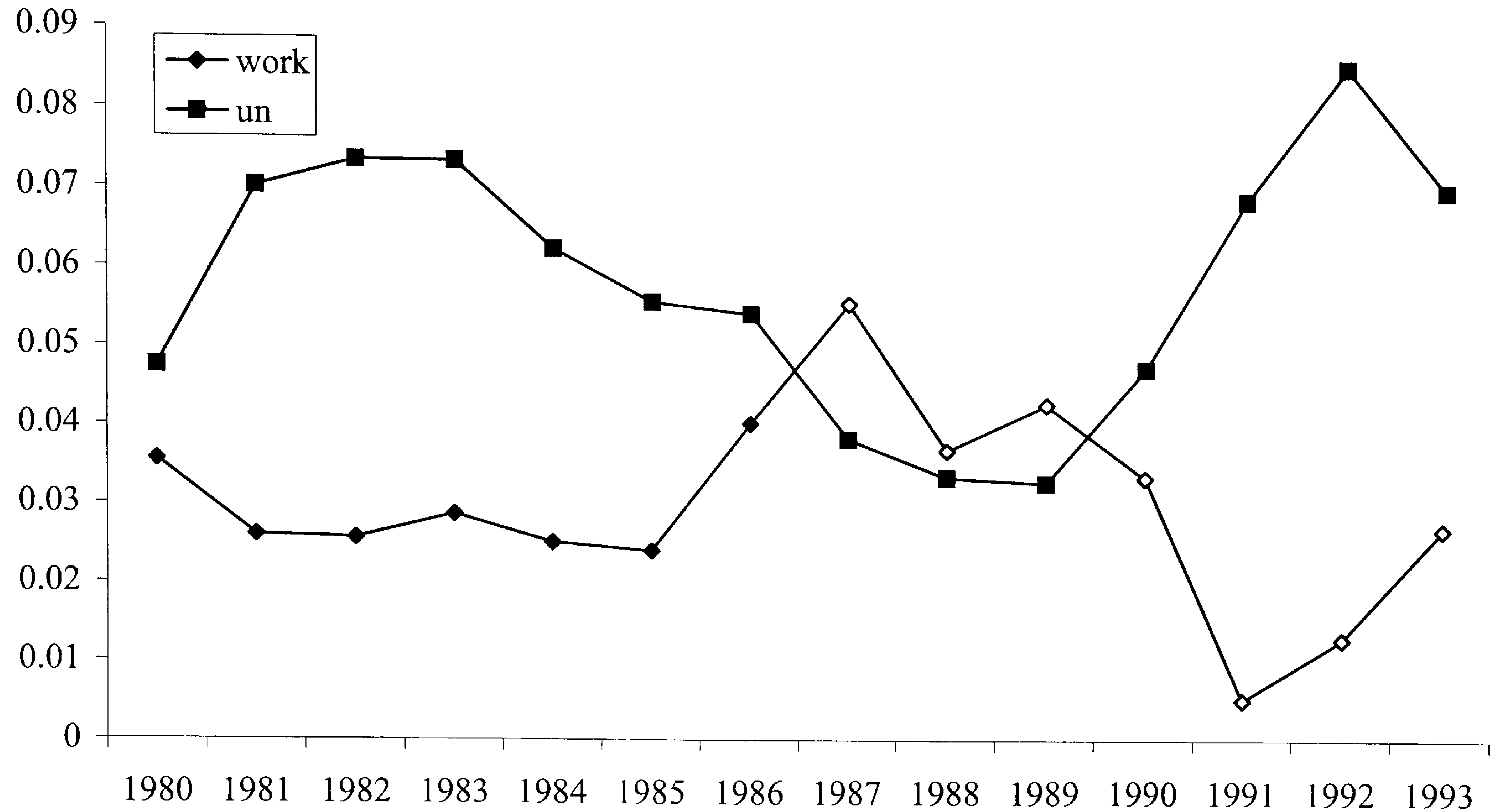


Figure 3.13 Social class effects (SCLOW)

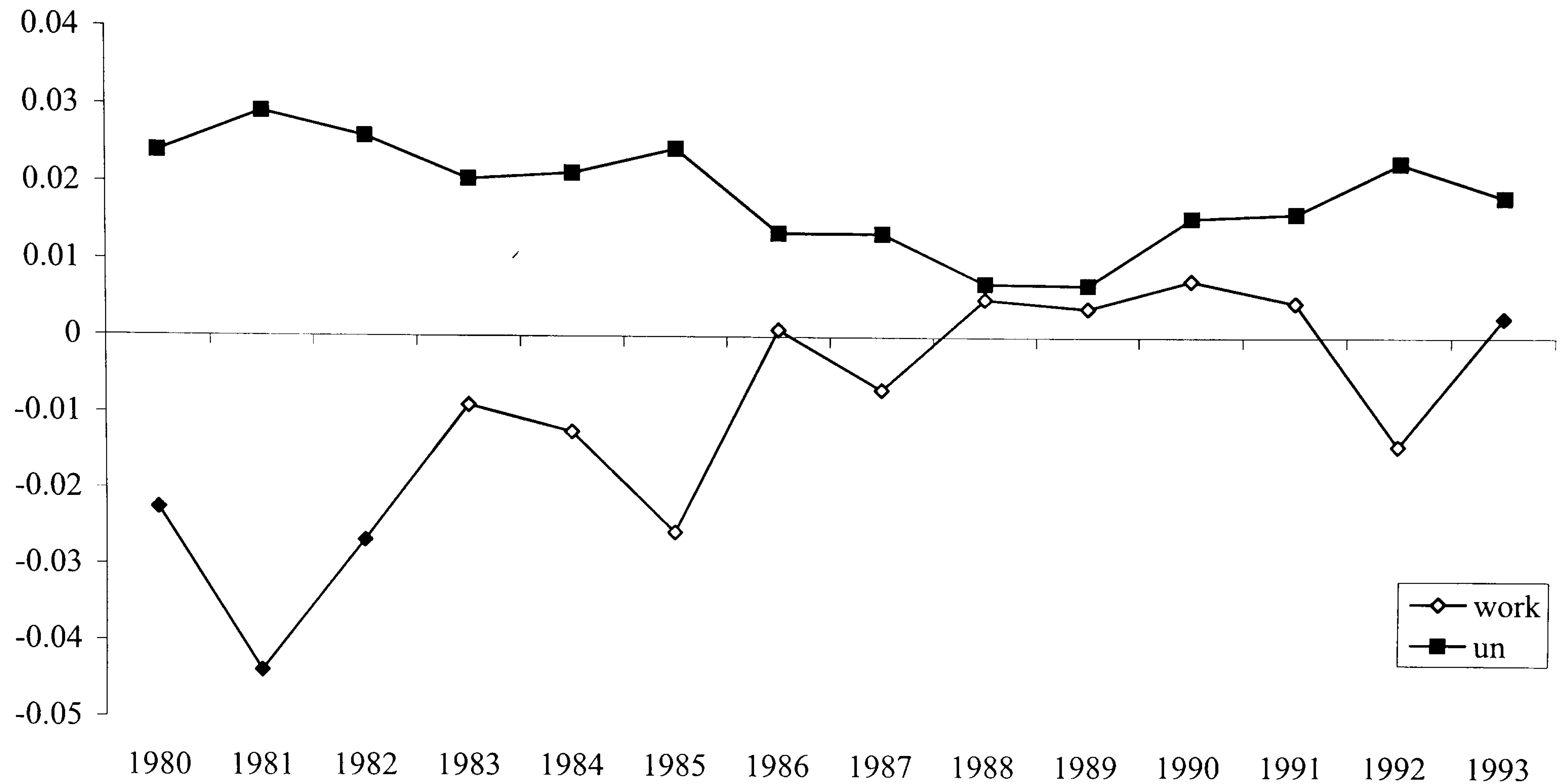


Figure 3.14 Independent school effects (INDEP)

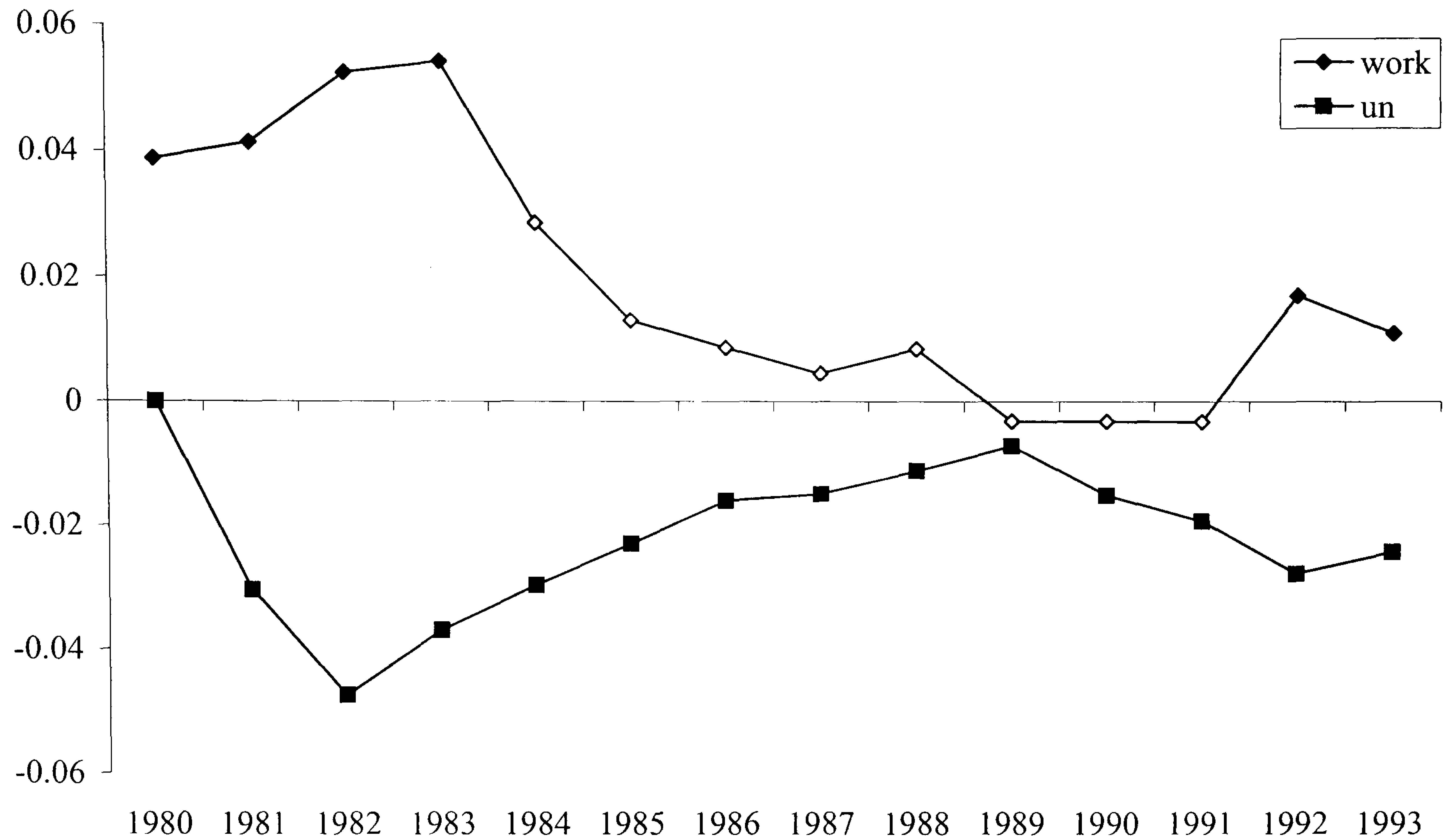


Table 3.4 Degree course effects: 1980 and 1993 graduate cohorts

	1980				1993			
course	WORK	STUDY	UN	TRAIN	WORK	STUDY	UN	TRAIN
BIOL	-0.006*	0.118*	-0.017*	-0.085*	-0.029*	0.129*	-0.030*	-0.074*
CHEM	0.002*	0.151*	-0.039*	-0.097	-0.128*	0.260*	-0.044	-0.067
PHYS	0.090*	0.093*	-0.061*	-0.106	-0.119*	0.208*	-0.017*	-0.055
MATHS	0.153*	0.011*	-0.064*	-0.082	0.029	0.024*	-0.033*	-0.007
COMP	0.257*	-0.028*	-0.059*	-0.149	0.186*	-0.018*	-0.022*	-0.115*
ECON	0.202*	-0.021*	-0.048*	-0.127*	0.148*	-0.027*	-0.028*	-0.088*
SOCIO	0.080*	-0.012	-0.009*	-0.054	0.083*	-0.039	0.000*	-0.048*
POL	0.063*	-0.014	0.013*	0.000	0.037*	0.012*	-0.006*	-0.034
LAW	-0.203*	-0.046*	-0.079*	0.343*	-0.282*	-0.064*	-0.071*	0.416*
OSOSCI	0.069*	-0.006	-0.022	-0.043*	0.055*	0.010*	-0.034	-0.037*
CLAS	-0.016	0.001	0.006	0.011	-0.005	-0.013*	0.001	0.014
ARTS	-0.101*	0.013	-0.016*	0.106*	0.037	-0.018*	-0.033*	0.020
MEUL	-0.001*	-0.026*	-0.028*	0.061*	0.013*	-0.043*	-0.018*	0.061*
ALMED	0.253*	-0.001*	-0.080*	-0.149	0.241*	-0.013*	-0.072*	-0.123*
ENGIN	0.307*	-0.024*	-0.079*	-0.182*	0.158*	0.013*	-0.033*	-0.116*
BUS	0.296*	-0.052	-0.078*	-0.158*	0.234*	-0.096*	-0.043*	-0.091*
EDU	0.186*	-0.035	-0.052	-0.087	0.229*	-0.083	-0.064	-0.070*
N	45584				56167			
LL	-12190				-12461			
pseudo R^2	0.15				0.10			

* denotes statistical significance at 5% level

Figure 3.15 Degree subject effects on P(UN) (reference group=HUM)

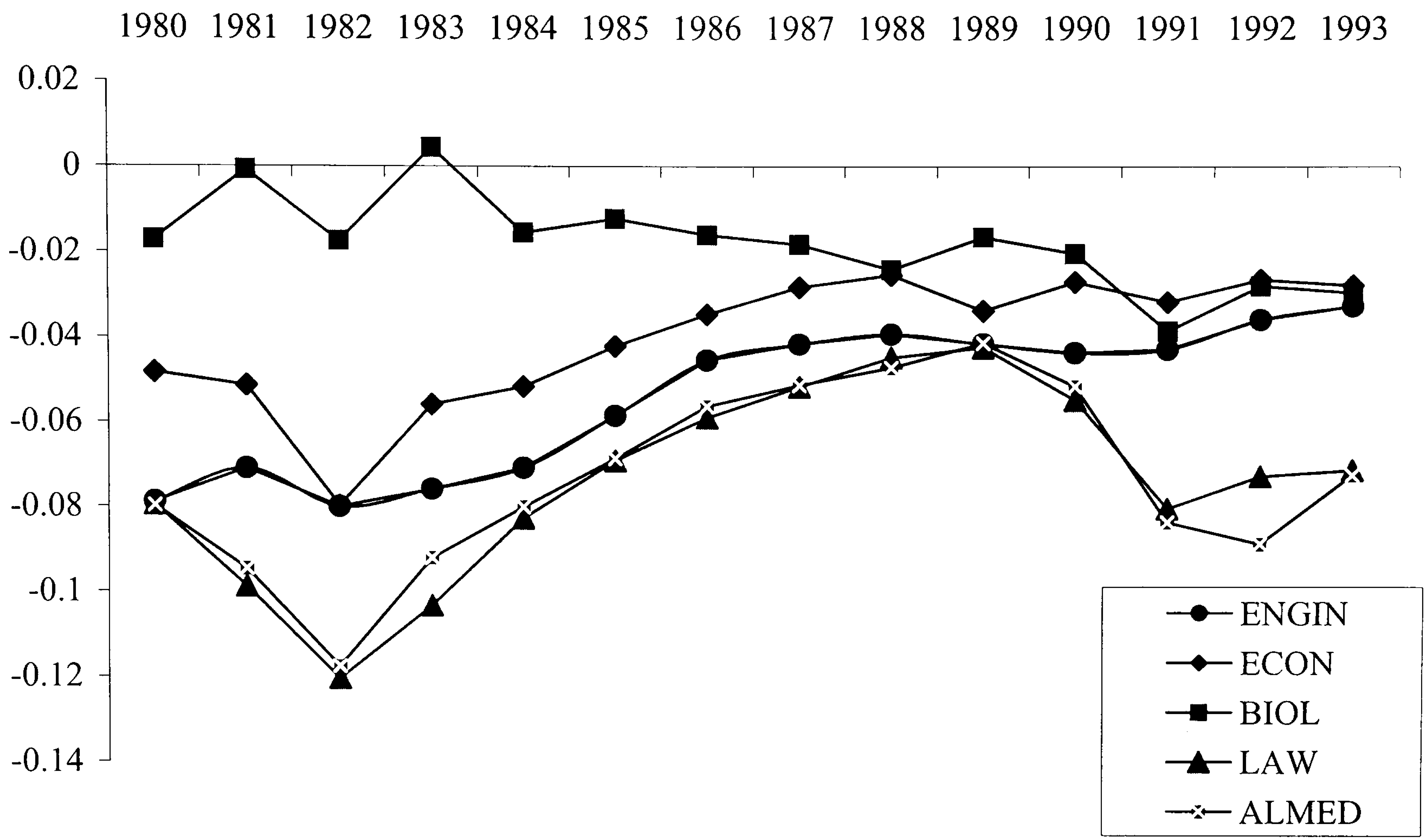


Table 3.5 'Oaxaca-Blinder' decompositions: gender differences in P(WORK) and P(UN)

	WORK			UN						
year	total Δ	% explained		total Δ	% explained		pseudo R^2		N	
		(M)[a]	(F)[b]		(M)	(F)	M	F	M	F
1980	12.6	63.3	75.0	2.0	38.2	29.0	0.17	0.11	27693	17891
1981	12.4	57.7	66.0	3.0	32.8	16.3	0.17	0.11	27975	18987
1982	9.9	70.8	72.8	3.6	31.5	21.1	0.18	0.11	28321	20234
1983	9.8	73.7	75.6	2.2	38.3	29.8	0.18	0.12	29365	21321
1984	9.2	75.1	80.7	2.2	37.2	29.1	0.19	0.13	28864	21112
1985	8.1	76.7	81.4	1.2	41.1	34.9	0.20	0.13	27541	21201
1986	5.6	75.9	94.1	2.1	31.5	22.9	0.18	0.13	27198	20761
1987	5.6	75.5	98.5	2.0	29.9	26.8	0.17	0.13	27275	21123
1988	5.3	87.9	89.4	1.6	32.7	28.4	0.16	0.13	27267	21237
1989	4.7	92.1	92.3	1.1	35.2	31.6	0.16	0.12	26832	21296
1990	3.5	90.7	83.0	2.4	25.5	22.6	0.14	0.12	26991	21945
1991	-0.1	48.6	49.2	4.5	4.4	2.5	0.12	0.11	27150	22621
1992	0.3	54.8	53.1	4.9	3.4	0.5	0.12	0.10	29167	24592
1993	-1.5	37.3	34.0	5.8	3.1	7.3	0.11	0.09	30134	26033

(a) Male coefficients are used as standard.
(b) Female coefficients are used as standard.

Table 3.6 'Oaxaca-Blinder' decompositions: other first destination outcomes

year	TRAIN			STUDY			OLF		
	total Δ	(M)	(F)	total Δ	(M)	(F)	total Δ	(M)	(F)
1990	-8.8	35.6	55.6	4.7	15.3	28.2	-1.6	13.5	8.8
1991	-9.5	32.3	48.2	6.0	25.4	26.9	-0.8	1.6	20.7
1992	-9.8	30.3	42.5	5.6	30.3	31.7	-1.1	33.7	30.8
1993	-9.4	37.2	40.3	5.8	22.7	34.6	-0.8	39.5	43.8

Table 3.7 Labour market 'specialisation' of degree courses (%)

	1988				1990				1992			
subjects	Modal sector[a]		Modal occupation[b]		Modal sector		Modal occupation		Modal sector		Modal occupation	
ENGIN	EC	51.3	EC	31.5	EC	47.6	EC	28.9	EC	39.4	EC	26.3
COMP	EC	30.2	CP	44.7	OS	31.3	CP	33.6	OS	28.6	CP	27.2
PHYS	EC	26.6	RD	20.7	EC	23.7	RD	19.7	PHE	23.6	RD	16.3
CHEM	OM	39.5	RD	27.5	OM	35.3	RD	27.2	OM	32.7	RD	25.0
BUS	BFI	47.2	ACC	40.6	BFI	45.3	ACC	38.7	BFI	40.8	ACC	32.8
ECON	BFI	53.8	ACC	38.6	BFI	51.4	ACC	38.8	BFI	43.7	ACC	31.0
ALMED	PHE	42.4	PHA	41.2	PHE	48.1	PHA	41.6	PHE	45.0	PHA	41.4
EDU	PHE	84.2	TP	44.9	PHE	86.1	TP	42.8	PHE	87.0	TP	34.5
MATHS	BFI	43.6	ACC	27.5	BFI	40.1	ACC	26.1	BFI	35.8	ACC	22.5

[a]Sectors: EC=Engineering & Construction; BFI=Banking, Finance, & Insurance; PHE=Public Administration, Health, & Education; OS=Other Services (incl. Consulting); OM=Other Manufacturing (incl. Chemical);
[b]Occupations: RD=Research and Development; CP=Computer Programming; TP=Teaching Primary schools (including nursery); PHA=Pharmacy; ACC=Accounting.

Table 3.8 Gender differences by modal sector of employment (%)

	1988			1990			1992		
sector	F	M	Δ	F	M	Δ	F	M	Δ
SE[a]	1.0	1.6	-0.6	1.0	2.0	-1.0	1.2	2.0	-0.8
PHE	31.0	16.6	14.4	34.1	18.5	15.6	35.4	22.3	13.1
OM	11.7	15.0	-3.4	11.4	14.5	-3.1	9.3	11.9	-2.7
EC	5.8	20.5	-14.6	5.8	19.3	-13.5	4.3	13.5	-9.2
BFI	17.7	22.6	-4.9	16.2	21.3	-5.0	13.2	18.1	-5.0
OS	5.1	7.3	-2.2	5.5	8.0	-2.6	5.5	8.1	-2.6
COMM[b]	24.6	14.5	10.1	22.8	14.2	8.6	27.4	21.2	6.3

[a] SE = Self-employment; [b] COMM = Commercial & Allied Services

Figure 3.16 Proportion of graduates in 'non-graduate' occupations (%)

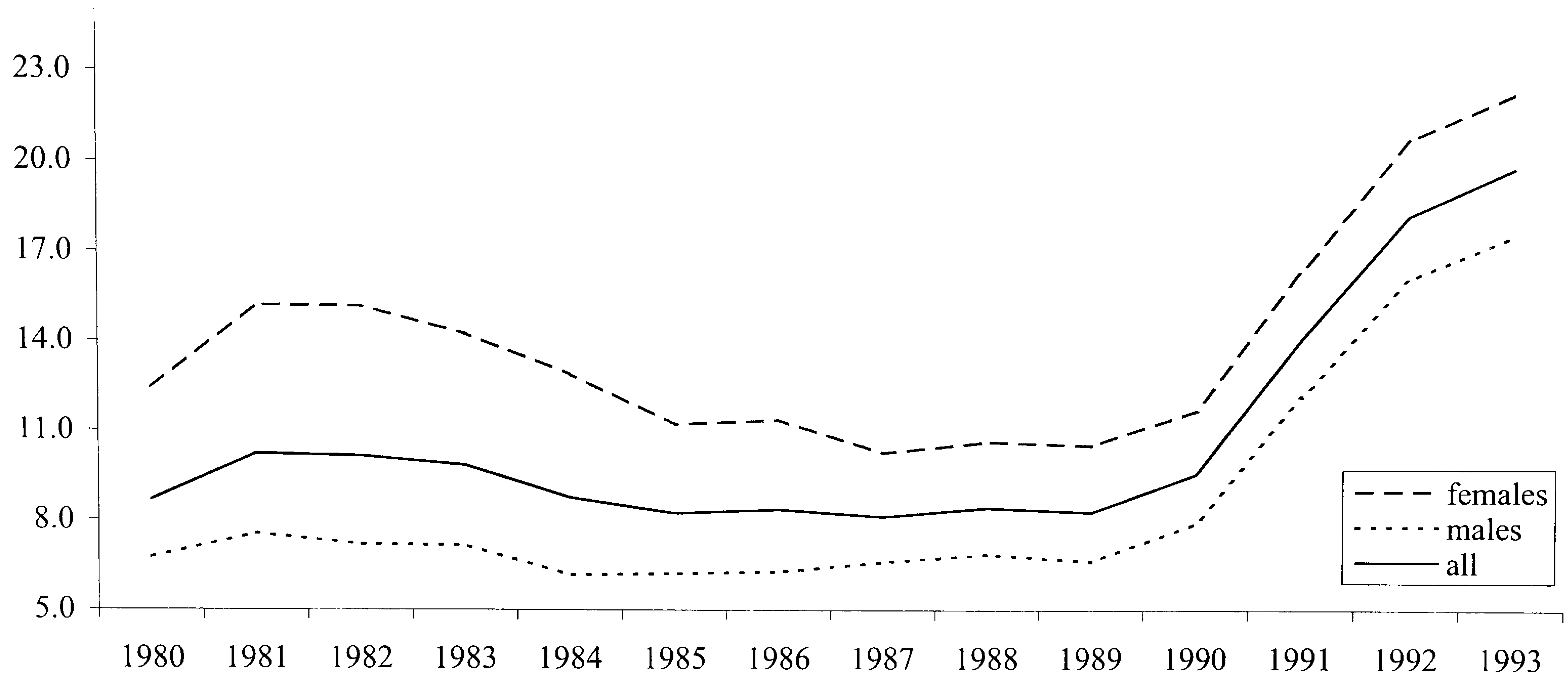


Table 3.9 Probit estimates: marginal effects on the probability to enter a 'non-graduate' occupation

	1983			1988			1993		
	Model			Model			Model		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
MANU		1.5 *	-14.8		1.2	-7.3		5.6 *	-23.5 *
SERV		4.2 *	-5.3		3.6 *	-20.3 *		15.6 *	-9.0
BIOL	-1.5 *	-1.0	0.1	-2.4 *	-1.9 *	-3.8	-3.5 *	-1.5	1.1
CHEM	-3.7 *	-3.5 *	1.4	-3.9 *	-3.7 *	-4.1	-9.1 *	-8.0 *	0.1
PHYS	-5.3 *	-5.1 *	-1.2	-4.1 *	-3.9 *	-4.4	-7.7 *	-6.6 *	-4.1
MATHS	-4.8 *	-4.8 *	-2.9	-4.7 *	-4.7 *	-3.6 *	-7.9 *	-8.8 *	-5.9 *
COMP	-6.4 *	-6.3 *	-4.9	-5.7 *	-5.6 *	-5.4 *	-15.9 *	-15.6 *	-13.7 *
ECON	-5.0 *	-5.2 *	-4.2 *	-4.1 *	-4.2 *	-2.3	-9.4 *	-10.4 *	-7.4 *
SOCIO	-1.8 *	-1.3	-2.7	-1.0	-0.7	-2.5 *	-5.5 *	-4.5 *	-6.5 *
POL	0.0	-0.2	0.1	-0.6	-0.7	-0.3	-3.4 *	-3.3 *	-3.2 *
LAW	-3.1 *	-3.3 *	-2.7 *	-2.9 *	-3.1 *	-0.9	-1.7	-2.6	-0.4
ARTS	-0.6	-0.9	-0.1	0.2	-0.2	2.6	-5.3 *	-6.7 *	-5.3 *
MEUL	0.2	0.0	0.8	-1.2	-1.2 *	-1.1	-3.4 *	-3.8 *	-2.7
ALMED	-6.3 *	-6.1 *	-6.4 *	-5.5 *	-5.3 *	-5.7 *	-17.0 *	-16.2 *	-16.7 *
ENGIN	-7.2 *	-6.8 *	-0.7	-6.8 *	-6.5 *	-7.1	-13.7 *	-12.1 *	-2.3
BUS	-6.2 *	-6.3 *	-5.1 *	-5.2 *	-5.2 *	-3.7 *	-12.1 *	-12.9 *	-8.8 *
EDU	-5.6 *	-5.1 *	-6.1 *	-4.5 *	-4.1 *	-5.4 *	-15.9 *	-14.0 *	-16.2 *
MALE	-1.7 *	-1.7 *	0.0	-0.1	-0.1	0.3	-1.7 *	-1.9 *	0.1
LEADUNI	-4.2 *	-4.1 *	-4.2 *	-1.2	-1.1 *	-1.2	-6.4 *	-6.0 *	-6.8 *
MATURE	0.2	0.6	-1.1	-0.4	0.0	-1.9	-6.9 *	-5.7 *	-8.3 *
POORDEG	2.4 *	2.2 *	2.0 *	2.6 *	2.5 *	2.9 *	8.4 *	7.6 *	8.2 *
SCLOW	1.1 *	1.2 *	1.1 *	0.2	0.4	-0.1	2.9 *	3.3 *	3.4 *
INDEP	-2.4 *	-2.7 *	-2.6 *	-1.2 *	-1.4 *	-0.5	-3.4 *	-4.2 *	-3.8 *
N	26569	26569	26569	28181	28181	28181	25715	25715	25715
LL	-7146	-7076	-7146	-6810	-6742	-6808	-11581	-11176	-11577
pseudo R2	0.12	0.13	0.12	0.11	0.11	0.11	0.1	0.13	0.1

(a) without controlling for student sector of employment

(b) controlling for student sector of employment

(c) using parental sector of employment as an instrument for employment sector of student

* denotes statistical significance at 5% level

Appendix 3A Likelihood ratio tests for the equality of MNL parameter estimates over time

The log likelihood function for the MNL model has the following form:

$$\ln \boldsymbol{L}_t = \sum_{i=1}^{N_t}\sum_{j=1}^{J}\left\{y_{ijt} \ln \boldsymbol{F}\left(\boldsymbol{X}_{it}\beta_j(t)\right)\right\}$$

where N_t is the number of graduates, J (= 4) is the number of possible first destination outcomes and t the graduate cohort to which the data refer.
The null and alternative hypotheses are:

H_0: $\beta(t)=\beta(t-1)$ j= 1,.....,4
H_1: $\beta(t)\neq\beta(t-1)$ j= 1,.....,4

Inter-cohort equality restrictions are imposed in turn on a subset of parameters of interest, leaving the remaining elements of the vector $\boldsymbol{\beta}$ unconstrained between consecutive years. For instance, to test whether LEADUNI effects have significantly changed between 1980 and 1981, we constrain the LEADUNI coefficient in the 1981 regression to be equal to the corresponding value already estimated for the 1980 cohort, while the other elements of $\boldsymbol{\beta}$ (1981) are left unrestricted.

Appendix 3B Secondary results

Table B.3.1 Multinomial logit estimates: 1981 graduates (TRAIN is default)

	WORK		STUDY		UN		OLF	
variable	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
BIOL	0.642	*0.071*	1.896	*0.098*	0.682	*0.086*	0.453	*0.166*
OBIOL	0.579	*0.074*	1.391	*0.108*	0.697	*0.089*	0.701	*0.155*
CHEM	0.694	*0.091*	2.183	*0.116*	0.453	*0.111*	0.151	*0.236*
PHYS	0.819	*0.097*	1.999	*0.126*	0.384	*0.123*	-0.140	*0.276*
OPHYS	0.984	*0.086*	1.690	*0.120*	0.518	*0.109*	0.798	*0.186*
MATHS	0.495	*0.077*	0.724	*0.120*	-0.222	*0.109*	-0.457	*0.229*
COMP	2.234	*0.203*	1.463	*0.264*	1.075	*0.239*	0.627	*0.501*
SOCIO	1.441	*0.088*	0.847	*0.146*	0.654	*0.110*	1.138	*0.175*
POL	0.619	*0.093*	0.361	*0.167*	0.434	*0.117*	0.541	*0.205*
LAW	0.788	*0.107*	0.750	*0.168*	0.546	*0.128*	1.117	*0.197*
OSOSCI	-1.710	*0.063*	-2.441	*0.164*	-2.571	*0.118*	-1.399	*0.164*
CLAS	0.453	*0.071*	0.237	*0.130*	0.229	*0.091*	0.517	*0.157*
ART	-0.155	*0.055*	-0.007	*0.097*	-0.151	*0.072*	-0.119	*0.132*
ALMED	-0.505	*0.091*	-0.200	*0.166*	-0.767	*0.136*	-0.644	*0.256*
MEUL	-0.306	*0.060*	-0.808	*0.126*	-0.665	*0.087*	-0.061	*0.145*
ENGIN	2.392	*0.126*	2.243	*0.160*	0.731	*0.168*	0.413	*0.337*
BUS	2.520	*0.104*	1.852	*0.136*	1.352	*0.121*	1.258	*0.204*
HUM	2.406	*0.114*	0.468	*0.227*	0.897	*0.142*	1.745	*0.200*
OTHER	1.479	*0.086*	1.610	*0.122*	0.842	*0.107*	1.137	*0.174*
POORDEG	-0.163	*0.029*	-1.965	*0.047*	0.404	*0.041*	0.051	*0.068*
SCLOW	-0.175	*0.036*	-0.075	*0.054*	0.113	*0.046*	-0.310	*0.091*
SC IIINM+SC OTH	-0.098	*0.035*	-0.034	*0.054*	0.133	*0.046*	-0.123	*0.083*
NOSCH	-0.021	*0.045*	0.073	*0.068*	-0.099	*0.060*	-0.052	*0.109*
INDEP	0.157	*0.046*	-0.008	*0.073*	-0.173	*0.063*	0.417	*0.098*
SCOREC	-0.088	*0.048*	-0.220	*0.068*	0.149	*0.072*	0.084	*0.117*
SCOREB	-0.158	*0.049*	-0.323	*0.072*	0.170	*0.072*	-0.024	*0.119*
SCOREA	-0.160	*0.049*	-0.241	*0.069*	0.239	*0.071*	-0.054	*0.117*
ALMATH	-0.318	*0.067*	-0.279	*0.103*	-0.133	*0.087*	-0.424	*0.152*
ALGENS	-0.176	*0.057*	-0.138	*0.085*	0.007	*0.073*	-0.199	*0.118*
ALCOUNT	-0.172	*0.056*	-0.166	*0.085*	-0.150	*0.075*	-0.264	*0.121*
HCOUNT	0.420	*0.039*	0.403	*0.059*	0.193	*0.051*	0.037	*0.092*
RES0	0.221	*0.037*	0.084	*0.055*	-0.054	*0.049*	0.020	*0.084*
RES1	0.022	*0.050*	0.014	*0.077*	-0.058	*0.064*	-0.181	*0.120*
RES2	0.019	*0.026*	0.098	*0.039*	0.098	*0.034*	0.064	*0.060*
RES3	-0.029	*0.026*	0.055	*0.043*	-0.005	*0.036*	0.010	*0.070*
RES4	-0.301	*0.164*	0.532	*0.212*	-0.442	*0.233*	0.119	*0.300*
RES5	-0.696	*0.126*	-0.466	*0.212*	-0.430	*0.175*	-0.795	*0.332*
RES6	-0.829	*0.066*	-0.465	*0.101*	-0.723	*0.089*	-0.723	*0.155*
RES7	-0.154	*0.062*	0.062	*0.093*	-0.034	*0.080*	-0.241	*0.134*
RES8	-0.533	*0.055*	-0.222	*0.084*	-0.253	*0.072*	-0.530	*0.125*
RES10	-0.454	*0.054*	-0.248	*0.081*	-0.250	*0.070*	-0.669	*0.128*
CDUR	-0.388	*0.060*	-0.192	*0.092*	-0.242	*0.079*	-0.474	*0.137*
MALE	0.762	*0.029*	1.168	*0.047*	1.040	*0.039*	0.334	*0.066*
LEADUNI	0.210	*0.048*	0.582	*0.070*	-0.346	*0.074*	0.141	*0.112*
MARRIED	-0.009	*0.083*	-0.238	*0.134*	-0.435	*0.125*	1.347	*0.146*
MATURE	0.061	*0.070*	0.642	*0.097*	0.245	*0.087*	-0.047	*0.154*
constant	0.396	*0.106*	-1.260	*0.164*	-1.557	*0.143*	-2.259	*0.248*
N	46962							
LL	-49409.1							
pseudo R2	0.15							

Table B.3.2 Multinomial logit estimates: 1987 graduates (TRAIN is default)

	WORK		STUDY		UN		OLF	
variable	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
BIOL	0.785	*0.086*	2.117	*0.111*	0.546	*0.121*	1.012	*0.144*
OBIOL	0.413	*0.091*	1.162	*0.124*	0.399	*0.130*	0.888	*0.151*
CHEM	0.565	*0.101*	2.178	*0.126*	0.091	*0.145*	0.129	*0.206*
PHYS	0.595	*0.103*	1.862	*0.129*	0.196	*0.145*	0.709	*0.180*
OPHYS	0.837	*0.106*	1.845	*0.133*	0.650	*0.140*	1.119	*0.168*
MATHS	0.298	*0.082*	0.506	*0.122*	-0.614	*0.136*	-0.388	*0.181*
COMP	2.505	*0.232*	1.983	*0.266*	1.495	*0.267*	1.516	*0.334*
SOCIO	1.160	*0.092*	0.648	*0.139*	0.429	*0.130*	0.757	*0.160*
POL	0.507	*0.097*	0.237	*0.161*	0.721	*0.129*	0.618	*0.173*
LAW	0.355	*0.099*	0.158	*0.158*	0.243	*0.138*	0.228	*0.187*
OSOSCI	-2.450	*0.068*	-2.500	*0.147*	-3.225	*0.169*	-1.875	*0.151*
CLAS	0.427	*0.080*	0.445	*0.128*	0.056	*0.124*	0.399	*0.148*
ART	-0.221	*0.061*	-0.008	*0.102*	-0.140	*0.096*	0.005	*0.120*
ALMED	-0.742	*0.089*	-0.583	*0.170*	-1.078	*0.169*	-0.770	*0.208*
MEUL	-0.203	*0.072*	-0.773	*0.141*	-0.256	*0.119*	-0.484	*0.156*
ENGIN	2.035	*0.142*	1.638	*0.183*	0.322	*0.226*	0.653	*0.270*
BUS	2.124	*0.114*	1.630	*0.145*	0.968	*0.147*	1.384	*0.179*
HUM	2.110	*0.133*	0.399	*0.220*	0.785	*0.179*	1.715	*0.187*
EDU	1.736	*0.192*	-0.135	*0.457*	1.125	*0.254*	1.562	*0.284*
OTHER	0.963	*0.088*	0.851	*0.129*	0.358	*0.130*	0.888	*0.153*
POORDEG	-0.046	*0.031*	-1.744	*0.049*	0.484	*0.048*	0.022	*0.057*
SCLOW	-0.053	*0.041*	-0.040	*0.057*	0.157	*0.059*	-0.382	*0.083*
SC IIINM+SC OTH	0.037	*0.040*	0.012	*0.057*	0.184	*0.059*	-0.151	*0.077*
NOSCH	0.021	*0.047*	-0.008	*0.066*	-0.057	*0.070*	-0.026	*0.089*
INDEP	-0.062	*0.041*	-0.222	*0.059*	-0.320	*0.064*	0.225	*0.071*
SCOREC	-0.090	*0.048*	-0.166	*0.065*	0.092	*0.079*	-0.043	*0.092*
SCOREB	-0.215	*0.050*	-0.184	*0.068*	0.074	*0.080*	-0.011	*0.093*
SCOREA	-0.175	*0.053*	-0.013	*0.071*	0.131	*0.083*	-0.113	*0.097*
ALMATH	0.258	*0.042*	0.200	*0.059*	0.068	*0.064*	0.095	*0.075*
ALGENS	0.033	*0.058*	0.017	*0.083*	-0.094	*0.084*	-0.031	*0.107*
ALCOUNT	0.013	*0.030*	0.108	*0.041*	0.068	*0.044*	-0.051	*0.056*
HCOUNT	-0.089	*0.029*	0.017	*0.046*	-0.055	*0.044*	-0.128	*0.067*
RES0	-0.547	*0.157*	0.550	*0.189*	-0.698	*0.274*	-0.316	*0.258*
RES1	-0.398	*0.142*	-0.136	*0.230*	0.219	*0.211*	-0.590	*0.330*
RES2	-0.694	*0.073*	-0.211	*0.103*	-0.305	*0.111*	-1.069	*0.159*
RES3	-0.139	*0.066*	-0.063	*0.093*	-0.118	*0.103*	-0.264	*0.116*
RES4	-0.346	*0.065*	-0.154	*0.091*	0.123	*0.095*	-0.483	*0.119*
RES5	-0.308	*0.063*	-0.159	*0.088*	-0.034	*0.095*	-0.487	*0.116*
RES6	-0.128	*0.069*	0.022	*0.095*	0.136	*0.102*	-0.349	*0.125*
RES7	-0.269	*0.076*	-0.112	*0.104*	0.094	*0.111*	-0.311	*0.134*
RES8	-0.149	*0.063*	-0.071	*0.086*	-0.114	*0.097*	-0.060	*0.103*
RES10	-0.119	*0.063*	-0.187	*0.088*	-0.170	*0.099*	-0.011	*0.103*
CDUR	0.310	*0.043*	0.020	*0.060*	-0.163	*0.066*	0.070	*0.079*
MALE	0.578	*0.032*	1.021	*0.047*	1.032	*0.049*	0.302	*0.059*
LEADUNI	-0.172	*0.054*	0.160	*0.070*	-0.114	*0.088*	-0.188	*0.100*
MARRIED	-0.394	*0.116*	-0.343	*0.170*	-0.302	*0.172*	0.234	*0.211*
MATURE	0.022	*0.078*	0.575	*0.107*	0.239	*0.109*	-0.123	*0.154*
constant	1.181	*0.123*	-1.012	*0.172*	-1.621	*0.186*	-1.101	*0.224*
N	48398							
LL	-46215.5							
pseudo R2	0.15							

Table B.3.3 Multinomial logit estimates: 1993 graduates (TRAIN is default)

	WORK		STUDY		UN		OLF	
variable	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
BIOL	0.683	*0.076*	1.502	*0.089*	0.425	*0.098*	0.806	*0.118*
OBIOL	0.769	*0.076*	0.779	*0.096*	0.451	*0.100*	0.834	*0.116*
CHEM	0.403	*0.090*	1.861	*0.099*	0.145	*0.113*	0.212	*0.153*
PHYS	0.272	*0.098*	1.547	*0.107*	0.331	*0.118*	0.142	*0.161*
OPHYS	0.670	*0.082*	1.202	*0.098*	0.422	*0.104*	0.869	*0.123*
MATHS	0.098	*0.071*	0.234	*0.092*	-0.300	*0.096*	-0.204	*0.126*
COMP	1.898	*0.143*	1.449	*0.161*	1.389	*0.159*	0.850	*0.216*
SOCIO	1.173	*0.080*	0.680	*0.104*	0.654	*0.101*	0.859	*0.121*
POL	0.565	*0.084*	0.034	*0.126*	0.432	*0.109*	0.499	*0.137*
LAW	0.342	*0.081*	0.377	*0.106*	0.223	*0.103*	0.086	*0.141*
OSOSCI	-2.120	*0.064*	-2.238	*0.106*	-2.463	*0.114*	-1.454	*0.111*
CLAS	0.404	*0.072*	0.398	*0.098*	-0.052	*0.101*	0.413	*0.117*
ART	-0.104	*0.055*	-0.215	*0.078*	-0.087	*0.075*	-0.038	*0.094*
ALMED	-0.078	*0.088*	-0.305	*0.134*	-0.494	*0.136*	-0.250	*0.165*
MEUL	-0.350	*0.068*	-0.812	*0.110*	-0.546	*0.099*	-0.631	*0.129*
ENGIN	2.225	*0.129*	1.753	*0.149*	0.839	*0.166*	1.051	*0.199*
BUS	1.609	*0.086*	1.472	*0.101*	1.016	*0.103*	0.870	*0.131*
HUM	1.341	*0.077*	-0.442	*0.135*	0.506	*0.104*	0.898	*0.122*
OTHER	0.918	*0.078*	0.862	*0.099*	0.303	*0.104*	0.671	*0.126*
POORDEG	-0.010	*0.028*	-1.117	*0.039*	0.542	*0.037*	0.117	*0.046*
SCLOW	0.033	*0.037*	-0.030	*0.047*	0.186	*0.048*	-0.165	*0.066*
SC IIINM+SC OTH	0.031	*0.035*	-0.022	*0.044*	0.121	*0.045*	-0.048	*0.058*
NOSCH	-0.098	*0.043*	-0.064	*0.053*	-0.088	*0.055*	-0.032	*0.071*
INDEP	-0.071	*0.033*	-0.186	*0.043*	-0.324	*0.045*	0.107	*0.054*
SCOREC	-0.108	*0.041*	-0.215	*0.052*	0.025	*0.058*	-0.225	*0.067*
SCOREB	-0.229	*0.043*	-0.272	*0.054*	0.005	*0.060*	-0.363	*0.071*
SCOREA	-0.244	*0.046*	-0.110	*0.057*	0.007	*0.063*	-0.417	*0.076*
ALMATH	0.232	*0.036*	0.217	*0.046*	0.094	*0.048*	0.112	*0.057*
ALGENS	0.085	*0.048*	0.085	*0.061*	0.073	*0.063*	0.053	*0.080*
ALCOUNT	-0.005	*0.026*	0.133	*0.032*	0.063	*0.034*	-0.064	*0.043*
HCOUNT	-0.085	*0.028*	0.049	*0.038*	-0.017	*0.038*	-0.162	*0.047*
RES0	-0.503	*0.139*	0.605	*0.157*	-0.879	*0.215*	-0.447	*0.220*
RES1	-0.407	*0.144*	-0.077	*0.192*	-0.291	*0.192*	0.004	*0.224*
RES2	-0.502	*0.064*	-0.338	*0.085*	-0.485	*0.087*	-0.610	*0.112*
RES3	-0.110	*0.057*	0.195	*0.071*	-0.067	*0.074*	-0.169	*0.089*
RES4	-0.352	*0.056*	-0.124	*0.071*	-0.212	*0.074*	-0.517	*0.094*
RES5	-0.284	*0.057*	-0.030	*0.071*	-0.132	*0.073*	-0.450	*0.094*
RES6	-0.245	*0.059*	-0.125	*0.075*	-0.199	*0.077*	-0.434	*0.097*
RES7	-0.270	*0.065*	-0.161	*0.083*	-0.239	*0.086*	-0.344	*0.107*
RES8	-0.044	*0.054*	-0.183	*0.070*	-0.022	*0.071*	-0.056	*0.082*
RES10	-0.180	*0.054*	-0.085	*0.068*	-0.183	*0.072*	-0.165	*0.084*
CDUR	0.542	*0.039*	-0.012	*0.050*	0.125	*0.052*	0.178	*0.063*
MALE	0.336	*0.028*	0.736	*0.036*	0.914	*0.038*	0.316	*0.046*
LEADUNI	-0.105	*0.051*	0.068	*0.060*	-0.335	*0.073*	-0.192	*0.082*
MARRIED	-0.136	*0.120*	-0.119	*0.151*	-0.346	*0.168*	0.406	*0.178*
MATURE	0.116	*0.065*	0.575	*0.078*	0.424	*0.080*	0.115	*0.107*
constant	0.927	*0.107*	-0.617	*0.135*	-1.038	*0.142*	-0.656	*0.175*
N	56167							
LL	-66193.4							
pseudo R2	0.10							

Chapter 4

Employment-related performance indicators for higher education institutions and data aggregation bias: evidence from the USR

4.1 Introduction

In recent years the practice of assessing the performance of higher education institutions with respect to a wide range of criteria has become considerably more widespread in the UK. This phenomenon has been observed across the whole UK public sector, including schools, the National Health Service, nationalised industries and local authorities (Smith, 1990). Broadly speaking, the development of widely accepted benchmarks, against which the performance of public sector organisations can be judged, aims primarily at promoting greater transparency in the way these organisations operate and at making them more accountable to the taxpayer. In particular, the use of university performance indicators as a monitoring tool for central government and funding councils over the range of educational services delivered by each institution in return for the funds received appears increasingly justified by the growing amount of public money allocated to the sector.[1] This concept was clearly stated by the Chancellor of the Exchequer in his 1999 pre-Budget Report: '[Given the substantial public investment in university students, it is particularly important that they are employable upon graduation. Better information is crucial to this aim. Work is already in progress to develop performance indicators, including those on employment outcomes, that will better inform the choices of perspective students]'.

In a typically regulated sector like education, performance indicators are also expected to serve other important purposes. The publication of suitable measures of performance on student progression, learning outcomes, teaching quality, research excellence and graduates' employability should encourage comparisons

---

[1] More than £8 billions of annual Government expenditures are allocated to higher education (Dearing, 1997).

between similar institutions and stimulate action to emulate best practice. More importantly, in the absence of price competition performance indicators may become an essential source of information to prospective students making their choices, especially in times when the costs of going to university are rising.[2]

Unlike other dimensions of university performance such as teaching quality (QAA) and research excellence (RAE), official league tables ranking institutions according to their capacity to produce employable graduates are far from being well established.

Previous studies have highlighted how the construction of reliable performance measures is bristling with difficulties. The first problem is defining employment success. Methodological issues such as whether or not postgraduate study should be treated as a 'positive' labour market outcome, or whether a distinction should be made between 'graduate' and 'non-graduate' occupations to account for differences in the quality of employment, demonstrate that a single set of indicators may fall short of capturing the multi-faceted nature of employment-related university performance (Johnes and Taylor, 1990; Smith *et al.*, 2000). A second stumbling block, common to other measures of performance like student achievement, is to devise indicators that satisfactorily accommodate the main factors influencing employment success. Crude indicators, which fail to adjust for university differences in the 'quality' of the student intake or in the mix of courses taught, may be seriously flawed, as Goldstein and Spiegelhalter (1996) have

[2] After the introduction of a £1,000 (means-tested) flat rate university fee in 1998, a proposal to set universities free to decide their own tuition fees outside government control and funding restrictions is currently under scrutiny in the UK (Greenaway and Haynes, 2000). Arguably, if fee liberalisation were fully implemented, market prices will function as a signal for university quality, and the role of league tables as a screening mechanism for potential applicants will be probably weakened.

shown. A third difficulty concerns the incidence of non-response on the reliability of survey data on university leavers' first destinations (Johnes and Taylor, 1990). Self-reported returns are usually only received from about 80% of the eligible population. If non-response happens to be a symptom of difficulties to find employment, those institutions with a high percentage of non-respondents may not only be overrated, but also less inclined to obtain information on reluctant responders, thus creating a self-reinforcing pattern. The fact that the first set of official employment-related indicators of university performance was published two years after their release was first announced is a testament to the complexity of the task.[3]

Beside these limitations, there is a more general methodological point that has received surprisingly little attention in the empirical literature on university performance, namely the sensitivity of the indicators to the level of aggregation of the data used to calculate them. One fundamental difference between two of the most influential UK studies on employment-related university performance authored respectively by Johnes and Taylor (1990) and Smith *et al.* (2000) is that the first uses university-level information, while the latter is based on individualised student data. There is a sizeable theoretical literature (Stoker, 1984; Van Daal and Merkies, 1984; Lewbel, 1989; Deaton and Muellbauer, 1980; Richards and Ben-Akiva, 1975) on the pitfalls of aggregating individual decisions - in our case first destination choices - to predict macro behaviour: university

[3] The first set of these official indicators was published in April 2001 in a separate volume of the Higher Education Funding Council for England (HEFCE) Report n. 2001/21. The indicators were the result of a pilot study conducted by the Performance Indicators Steering Group (PISG) established in 1997 with a membership drawn from government departments, the funding councils and representative bodies. The study was based on a sample of 3000 graduates from the 1999-2000 academic year and it was set up to investigate the feasibility of central data collection for the FDS.

performance, in our context. Moreover, empirical research on transportation mode choices has clearly shown that the estimation of *macro* behaviour is highly sensitive to the way individual characteristics and decisions are aggregated (Talvitie, 1973; Koppelman, 1975a, 1975b, and 1976; Westin, 1974; Watson and Westin, 1975; Nam, 1997). Given that universities are understandably sensitive to their position in the tables, an analysis of whether performance indicators are robust to data aggregation appears warranted. Moreover, the relentless dissemination in the media of university rankings based exclusively on aggregate statistics bears witness to the fact that data considerations are either completely ignored or regarded as uninteresting. It must be stressed that the prescription of specific solutions on how reliable indicators of performance should be developed is beyond the scope of this chapter. Nonetheless, we believe that our analysis on aggregation issues can inform the process towards a satisfactory solution to the complex and ambitious task of producing widely accepted employment-related university league tables.

The rest of the chapter is structured as follows. Section 4.2 reviews some of the literature both on the construction of performance indicators and on the distortions associated with data aggregation. It attempts to clarify the nature of the aggregation problem and how it relates to the development of employment-related output measures in higher education. Section 4.3 explains the methodology used to construct university league tables both from aggregated and disaggregated data. Section 4.4 illustrates the main features of the sample. In Sections 4.5-4.7 we compare micro and macro outcomes, both in terms of parameter estimates and university rankings, using different modelling specifications with the aim to establish the existence and estimate the size of an aggregation bias. In particular, Section 4.7 presents a Monte Carlo simulation designed to single out the bias originating from data aggregation from other types of bias. In Section 4.8, we study the sensitivity of the bias to alternative aggregation procedures. Finally, Section 4.9 concludes the chapter summarising the main findings.

4.2 Literature review

The analysis developed in this chapter draws upon two distinct bodies of literature, which will be reviewed separately. The first, presented in Section 4.2.1, is specific to the construction of performance indicators for the education sector, in general, and to employment-related measures of university output, in particular. The second, discussed in Section 4.2.2, is concerned with the effects and measurement of the bias originating from data aggregation.

4.2.1 Performance indicators

Over the last two decades there has been an increasing interest in the construction of performance indicators in the UK targeting a broad range of public sector activities like education, health, local authorities, and other social services. This has coincided with the beginning of the systematic publication of comparative data as part of a conscious government policy aimed at introducing or enhancing accountability. Smith (1990) discusses the background to this interest and reviews the progress made in the development of performance indicators during the 1980s in England and Wales. Important examples are the introduction in 1981 of the code of practice for the publication of local authorities annual reports, which imposed the dissemination of a wide range of comparative statistics, the publication in 1983 of 'grey books' containing over 123 performance indicators for local health authorities, and the launch in 1987 of the *University Management Statistics and Performance Indicators* booklets, a package of 39 indicators dominated by expenditure data jointly produced by the Committee of Vice-Chancellors and Principals (CVCP) and the University Grants Committee (UGC).

Education is one of the areas where performance indicators have found a particularly fertile ground for development and use. The bulk of research on performance indicators for the education sector in the UK seems to have concentrated on assessing the contribution of schools to students' achievement. In an influential study on the construction of league tables for LEA schools based on public examinations' results, Goldstein and Spiegelhalter (1996) discuss the main statistical issues involved and the limitations that typically affect these measures of school output. The analysis highlights the inadequacy of using crude rankings that fail to make proper allowance for students' pre-existing achievements. In addition to that, even 'value added' tables properly adjusted for the impact of contextual factors like prior achievement, are likely to suffer from a large margin of error or 'uncertainty', especially when cell size, in this case the number of students in each school, is small.

Acknowledging the limitations of survey data, Johnes and Taylor (1990) wrote an influential book on performance indicators in higher education. Chapter 6 of the book is dedicated to the construction of league tables ranking universities on the basis of their success in producing employable graduates. Using aggregate USR statistics on students' first destinations between 1983 and 1986, the study produces two sets of adjusted indicators constructed from the raw percentage of graduates obtaining a permanent job and graduates proceeding to further education or training, respectively. The authors find that about 70% of the inter-university variation in graduates' first destinations is explained by the corresponding variation in subject mix.

Smith *et al.* (2000) construct employment-related performance indicators and the resulting university league tables using individualised USR data on 1993 university leavers in the UK. The probability that the student is either unemployed or inactive (negative outcome), rather than employed or in further study (positive outcome) six months after graduation is modelled and estimated using a binomial probit model. Individual university marginal effects on the probability of the negative outcome are derived after controlling for university differences in student personal and course-related characteristics. These effects are then used to rank institutions and compile university league tables of employment success. The low correlation between adjusted and unadjusted (based on raw unemployment proportions) tables confirmed the importance of controlling for the influence of key factors like prior qualifications, subject studied, and degree performance. Furthermore, as in Goldstein and Spiegelhalter (1996), the authors address the issue of the confidence in the rankings and find that in most cases university performance cannot be separated with a sufficient level of confidence.

In 2002 a set of employment-related performance indicators was included for the first time with other indicators in the main annual publication monitoring the performance of the higher education sector in the UK (HEFCE Report 2002/52), after being preceded by a pilot study published in a separate volume of the 2001 publication (HEFCE Report 2001/21). The 2002 Report is based on data submitted to HESA by 166 publicly funded institutions of higher education in the UK for the 1999-2000 and 2000-2001 academic years. The indicators are based on information about full-time, home-domiciled undergraduate students only. Two indicators of employment success are proposed with and without the inclusion of individuals in further study among the employed. For each indicator an actual value, a benchmark value and context statistics are provided. The benchmarks are constructed by estimating a random-effect multi-level model, using a wide range of adjustment factors such as age on entry, entry qualifications, subject of study, gender, ethnic group, social background, degree classification, and whether or not the student was on a sandwich course.[4] In addition to these, local factors (unemployment in the locality where the university is located) and institutional factors (average A-level scores) are also considered. The context statistics provided include population numbers on which the percentages are based, the FDS response rate for each institution and the proportion of the population who are excluded from the indicators.

[4] It was unclear whether to consider sandwich course as a factor outside or within an institution's control and, consequently, if it should be omitted or included in the benchmark. The decision to include the sandwich courses marker in the benchmark of the 2001 pilot study was queried by some of the institutions affected. Consequently, it was agreed to omit it in the 2002 publication. However, to allow for comparison with the 2001 year's figures, the 2002 publication reports the extent of the sandwich effect.

4.2.2 Data aggregation

The econometric literature on the problem of aggregation is far too vast and dispersed to be surveyed comprehensively in this section. Given the specific focus of the chapter, we restrict our attention to aggregation over individuals, as opposed to other types of aggregation (over commodities, regions, time). Furthermore, we concentrate on the effects of aggregation when the underlying micro relations are non-linear, as in the case of discrete choice models. Broadly speaking, theoretical research on aggregation has shown that if individual heterogeneity is not adequately accounted for in the aggregation process, aggregate behaviour will incorrectly reflect actual individual choices. A classical example of inconsistent aggregation, but yet often used in empirical research, is when macro variables are derived as mere averages of their micro counterparts in a *representative agent* (*RA* henceforth) fashion. This naïve aggregation technique consists of estimating the macro equations with variables representing the characteristics and choices of a fictitious average individual. Deaton and Muellbauer (1980) show that in a linear model of consumer choice with an additive error term, the *RA* approach, which regresses demand on average characteristics, can be used to obtain consistent estimates of the underlying individual demand curves.[5] However, when micro relations are non-linear, consistency is no longer assured (Van Daal and Merkies, 1984; Lewbel, 1992).

A number of studies have attempted to identify the conditions under which macro behaviour can be consistently estimated using aggregated data alone. The first approach consists of including additional information on the probability

[5] The effects of aggregation when the micro relations are linear were originally examined by Gorman (1953) and Theil (1954).

distribution of the aggregate explanatory variables - such as variances and covariances - in the aggregate functions to be estimated. When individual data are incomplete or unavailable, the solution frequently adopted is to use density functions which offer computational advantages (Kelejian, 1995; Stoker, 1982). A typical example is represented by functions of the exponential family (normal, log-normal, gamma, beta distributions). With discrete choice models, it has been shown that the binary probit model is the only case that lends itself to a tractable solution and requires the assumption of multinomial normally distributed explanatory variables (McFadden and Reid, 1975).[6] Lewbel (1992) considers log-linear models and illustrates the distributional assumptions required for consistent aggregation. Kelejian (1995) suggests a procedure for reducing the aggregation bias with logit models based on aggregated cell data. Allenby and Rossi (1991) examine the aggregation properties of nested logit models in a consumer-choice framework, and find that under specific conditions, estimation of aggregate models is theoretically justified. More recently, Van Garderen, Lee, and Pesaran (2000) have proposed an alternative approach to aggregation. The focus shifted from the issue of parameter consistency, to the objective of making optimal predictions of the aggregated model conditional on the macro information available. The aggregated dependent variable is derived as the (optimal) forecast, in the sense of least mean squared prediction errors, with respect to the aggregated information set. Again, this procedure requires *ad hoc* restrictions on the probability distribution and on the nature of the unknown micro relations.

---

6 Westin (1974) derives a similar procedure for the logit model with normally distributed variables. However, the moments of the resulting (S_B) distribution are analytically intractable and numerical integration is required to compute the moments of interest.

A sizeable bulk of empirical literature on the aggregation of discrete choice has taken place in the area of transportation mode choice (Talvitie, 1973; Koppelman, 1975a, 1975b and 1976; Westin, 1974; Watson and Westin, 1975; Nam, 1997). A recurrent issue discussed in these studies is the use of disaggregated models of travel decisions - such as trip mode (car, bus or train) or trip frequency - to predict aggregate travel demand. For planning and forecasting purposes some level of aggregation is not only inevitable, but it is also necessary to verify the validity of the relations between the macro variables and to test whether predictions based on individual choice models are transferable from one population to another. For example, the effect of introducing a commuter line into a new area can only be predicted by extrapolating models estimated in other areas where such a choice exists (Westin, 1974). In urban transportation studies, data collected by means of a home-interview survey, are typically aggregated, or averaged, geographically across individuals into traffic zones (such as districts, cities or regions) and/or across trip types (centre to centre, suburbs to centre, periphery to suburbs). All the travellers between any pair of zones are grouped together and individual characteristics are averaged for each group. Travel demand is then estimated with mean data, using the traffic zone as the observation unit.

A similar framework can be easily applied to first destination choices of university leavers. In general terms, individual student characteristics and first destination decisions are averaged across individuals up to the university level, generating a dataset in which institutions become the unit of observation. The research on transportation mode choices has clearly shown that the prediction of macro behaviour is highly sensitive to the way individual characteristics and decisions are aggregated. The ideal aggregation procedure would be to sum or average of each individual's predicted behaviour estimated with individual-level data. In this case individual heterogeneity is not affected by aggregation and micro behaviour is consistently represented at the macro level. However, when individual data are unavailable or incomplete, aggregated data is the only option, and measuring the aggregation bias becomes a by-product of any aggregation procedure. Given the restrictive informational requirements on the distribution parameters of the explanatory variables imposed by the analytic approaches presented earlier, empirical studies have used techniques based on weaker distributional assumptions - such as Taylor series expansions and population segmentation - to reduce the aggregation bias and improve the aggregate predictions generated by the *RA* approach. Broadly speaking, the improvement consists of accounting, at least to some degree, for individual heterogeneity. Details on how these methods are developed and applied to the construction of performance indicators are given in Section 4.8.

4.3 The construction of employment-related performance indicators

The analysis of graduates' first destinations presented in the previous chapter has shown that leavers' choices are influenced by a number of personal and university-related factors. We found significant effects associated with the subject of study. For instance, relative to Social Sciences, Science subjects increase the probability of postgraduate studies, whereas more vocational subjects such as Business Studies, Education, and Allied Medicine typically enhance the probability of employment. Personal characteristics also matter. Female graduates are more likely to enter employment, as are individuals from more affluent backgrounds, while students with lower A-level scores and/or with a poorer degree performance are more likely to experience unemployment. These effects need to be controlled for in order to disentangle the genuine contribution that each university makes to the early career success of its graduates. The analyses presented in the previous chapters of the Thesis have also shown that the university type that students attend can influence graduates' employability significantly, even after making allowances for student differences in personal or other institutional factors. In the following sections we will examine these university effects in greater detail by focusing on single institutions rather than university types.

Following Smith *et al.* (2000), we use graduate unemployment to measure the success of universities in producing employable graduates. Graduates are classified into two groups on the basis of their self-reported first destinations. The first group comprises the employed, the postgraduates, and the professional trainees, while the second group includes the unemployed and those out of the labour force (OLFU). The rationale behind this division is to separate 'positive' from 'negative' outcomes. This reflects the government's stance that graduates should achieve positive outcomes, given the amount of public resources allocated to higher education. We are aware that this 'black and white' division gives only a simplified and partial account of graduate employability. For instance, this chapter does not contemplate a distinction between further study and employment (Johnes and Taylor, 1990), nor, within the latter, between graduate and non-graduate employment (Smith *et al.*, 2000). However, given that the main purpose of the chapter is to investigate the effect of data aggregation on the employment-related university output measures, we choose to keep methodology simple. A finer categorisation of the dependent variable would benefit the analysis only marginally at the cost of making the construction of performance indicators computationally more demanding.

In the rest of the section, we separately discuss the modelling strategy used to construct performance indicators based on individualised data (Section 4.3.1) and university-level information (Section 4.3.2).

4.3.1 Performance indicators based on student-level data

The first step towards the construction of employment-related university performance indicators consists of estimating a binomial logistic model of the probability that an individual achieves a 'negative' outcome (this category will be referred to as OLFU hereafter). More formally,

$$p_i = \Pr[OLFU = 1] = \frac{\exp\left(X_i'\beta + Z_i'\gamma_j\right)}{1+\exp\left(X_i'\beta + Z_i'\gamma_j\right)} \qquad i=1,\dots\dots,n \qquad (4.1)$$

where p_i is the OLFU probability for the i^{th} individual, X_i is the vector of control variables including personal and educational characteristics, and Z_i is a vector of university dummies. After estimating Equation (4.1), university performance is measured by the institution-specific marginal effects ϖ_j. These are obtained by averaging (over individuals) the change in the individual OLFU probabilities associated with graduating from university *j* relative to university *0* (default). More formally,

$$\varpi_j = \frac{1}{n}\sum_{i=1}^{n}\left(\Pr[OLFU=1]_{z_j=1,x_i} - \Pr[OLFU=1]_{z_j=0,x_i}\right) \quad i=1, \;,n; j=1,..\; m \qquad (4.2)$$

For simplicity, the worst-performing university was omitted from the regression and used as the base institution to which these relative marginal effects all refer.[7]

4.3.2 Performance indicators based on university-level data

If individualised data are not available, the assessment of university performance has to rely on aggregated data. As a consequence, the OLFU probabilities have to be estimated directly at the university level. The primary purpose of this chapter is to compare performance indicators, and ultimately university league tables, constructed from *micro* and *macro* data, respectively. In order to ascertain the 'true' effect of data aggregation on the outcomes, it is important that comparisons are made on a 'like for like' basis. A first aspect of comparability pertains to the functional form of *micro* and *macro* equations. Given the non-linearity of Equation (4.1), we estimate by Weighted Least Squares (WLS) the following *macro* equation:

[7] With binary logistic models, rankings based on university coefficients are equivalent to rankings based on university marginal effects, as the order of magnitude is preserved.

$$\ln\left[\frac{P_j}{1-P_j}\right] = X'_j\beta + \varepsilon_j \qquad j=1,\dots, m \qquad (4.3)$$

with weights $w_j = \sqrt{n_j\pi_j(1-\pi_j)}$, where n_j is the student population at university j, $\pi_j = \frac{e^{X'_j\beta}}{1+e^{X'_j\beta}}$ is the predicted proportion of unemployed graduates from university j, $P_j = \frac{1}{n_j}\sum_{i=1}^{n_j} p_i$ is the proportion of leavers from the j^{th} university who are still unemployed six months after graduation, $X_j = \frac{1}{n_j}\sum_{i=1}^{n_j} X_i$ is a vector of mean values of selected controls for the j^{th} university, and $\varepsilon_j \sim N\left(0, \frac{1}{n_j P_j(1-P_j)}\right)$.[8]

The weights w_j, used to correct the heteroscedastic error term in Equation (4.3), depend on the unknown parameters β and a two-step procedure is required. The residuals $\hat{r}_j$ estimated from Equation (4.3) are then used as indicators of employment-related university performance.[9] In fact, residuals capture university differentials in leavers' unemployment, which remain 'unexplained' after controlling for differences across institutions in key factors both institution-related

[8] This model is very similar to a logit regression with grouped data (Greene, 2000, p. 894). The only difference is that our unit of observation is the university j rather than the individual i.

[9] There are some important differences between our approach and the one devised in Johnes and Taylor (1990): in our study the dependent variable is the university share of unemployed rather than the share of employed. In addition, the dependent variable is not adjusted for subject mix beforehand, but subject effects are included in the main regression with the other explanatory variables.

(university type, location, and subject mix) and student-related (academic ability, gender composition and socio-economic background). Thus, residuals are believed to reflect, at least to an extent, unobserved differences in graduate employability directly associated with the specific institution where students gained their degree. For the sake of comparability between the residuals $\hat{r}_j$ and university marginal effects discussed in the previous section, we note that the micro coefficients γ_j from which the marginal effects ϖ_j are derived can also be considered as residuals as they are computed conditional on the index $X'\beta$. In other words, the marginal effects ϖ_j are affected by changes in $X'\beta$ and encapsulate any residual effect which is omitted from $X'\beta$.

A second key aspect of comparability pertains to the model's specification, and in particular, to the correspondence between the vectors X_i and X_j used in Equations (4.1) and (4.3), respectively. This issue will be discussed in the next section.

4.3.3 The choice of control variables

Given the purpose of this chapter, the need to control for a rich set of 'contextual' factors to produce 'adjusted' measures of university performance has to be .balanced with the objective to ensure 'like-with-like' comparisons between *macro* and *micro* outcomes. It is often the case that information available at the micro level has no direct macro equivalent. For instance, with respect to the USR, important information on student socio-economic background[10] and pre-

[10] As pointed out in Chapter 2, social class information is not directly available from the USR micro files, but was obtained through matching USR codes on parental occupation with SOC codes.

university qualifications is not available at the university level. When micro data are available, like in our case, ensuring a direct correspondence between micro and macro variables does not impose *per se* particular limitations to the specification of Equations (4.1) and (4.3). In fact, any student-level variable can be averaged across individuals from the same institution to generate its university-level counterpart. For instance, the individual attribute of coming from a working class family becomes, in aggregate, the proportion of working class students in each university. However, a more binding constraint to the choice of X_j is represented by the limited sample size at the university level. The number of 'old' institutions surveyed in the USR is less than sixty. This imposes a restriction to the dimensions of the vector X_j, and consequently, on X_i. We focus on the following adjustment factors (see Table 4.1 for a full description):

i. *Subject studied (mix)*. Four broad subject areas are considered: Science (including Life, Physical and Mathematical Sciences), Social Sciences, Arts, and vocational-oriented courses. These categories are self-explanatory except for the last, which includes Engineering, Computer Science, Allied Medicine, Law, Education, and Business. The rationale for creating a separate category for vocational-oriented subjects is that institutions with a high proportion of students enrolled in these courses are expected to have a relatively lower proportion of unemployed or inactive graduates;

ii. *Academic ability*. Accounting for university differences in the 'quality' of the student intake is crucial to measure the added value of each institution to the employability of its graduates. We use A-level grades and degree performance as proxies for the academic ability of the graduate population.[11] The model also includes a control for A-level Mathematics as an indicator of specific skills (mathematical ability) that are typically valued in the labour market;

iii. *Entry qualifications.* We consider the distinction between A-level qualifications, which represent the main route into higher education in the UK, and other 'non-traditional' qualifications like BTEC diplomas and Advanced GNVQs;

iv. *Socio-economic background.* Social class based on parental occupation and type of school attended are used as proxies for graduates' socio-economic status. If students from more affluent backgrounds are connected to wider business networks, social class is another important contextual factor to take into account;

v. *Gender.* The analysis presented in the second chapter of the Thesis as well as results from other studies on first destination (McKnight, 1999) have shown that females, on average, are more likely than males to find a job in the first six months after graduation. This suggests that differences between universities in gender composition also need consideration;

vi. *Method of study.* We account for 'within' (micro-level) and 'between' (macro- level) university differences in the mode of study, namely full time, part-time, and sandwich. Intercalated degrees allow students to establish early ties with the labour market and accounting for inter-university differences in the proportion of graduates doing sandwich courses is arguably an important adjustment factor to assess employment-related university performance;

---

[11] As in Chapter 2, SCE Highers grade scores have been converted into an A-level equivalent (see Chapter 2, Appendix 2A for details).

vii. *Age.* We also control for 'within' and 'between' university differences in the age composition of the graduate population;

viii. *Labour market factors.* We consider indicators of the unemployment rate i) in the district or county where the university is located, and ii) in the district or county of prior residence of graduates, to account for regional labour market effects on graduates' first destinations.[12] It is therefore possible that the employability of graduates coming from high-employment districts or counties, or graduates who graduated from universities located in high-unemployment areas, is lower because of regional labour market factors that have little to do with the university attended.[13]

In the next section, we illustrate the composition of the sample and present descriptive statistics of the main variables.

4.4 Data and summary statistics

The analysis exploits USR student-level data on a complete cohort of students who left university in 1993 in the UK. The sample consists of 63,515 individuals from 55 different institutions. In addition to the selection criteria discussed in Section 2.5,[14] we have excluded approximately 200 individuals from two minor

---

[12] The information is based on the 1991 census. The source is www.census.ac.uk/casweb.

[13] It could be argued that to some extent good universities are able to foster economic activity at their doorsteps and, therefore, regional effects are not entirely independent of university performance.

[14] Non-respondents are typically excluded from all the studies reviewed in Section 4.2.1. Differences between institutions in non-response rates could bias employment-related university performance if unaccounted for. For instance, if non-respondents tend to be unsuccessful in obtaining a satisfactory destination, then the performance of those institutions with relatively high

institutions due both to cell size considerations and to the unusual composition and first destination patterns of their graduate population.

Table 4.2 shows summary statistics of the data. The figures reveal that 16.4% of graduates are still unemployed or non-employed six months after graduation, and this proportion ranges from 7.6% to 29.4% across universities. About 31% of the individuals with A-level qualifications achieved at least 26 points in their best three A-level passes (ABB+).[15] However, there is a remarkable variation between institutions. Controlling for these differences in the academic ability of the student intake is important to single out 'residual' university effects. The average proportion of students with at least upper second class degree (GOODEG) is around 59%, but it is considerably higher for some institutions (up to over 77%). The average sample proportion of students from working class homes (SCLOW) is 17% and ranges between 7.7% and 24.3%. We also note that 23.5% of the graduates in the sample were educated at an independent school (INDEP). This average figure conceals considerable variation between universities with peaks as high as 50% at the upper end, and institutions with hardly any former independent school pupils, at the lower end of the distribution. The proportion of students in vocational-oriented courses (VOC) varies remarkably across institutions (from a

---

non-response rates may be biased upwards. It is also possible that the proportion of 'unknowns' is commensurate to the zeal that individual universities put into pursuing non-respondents. If this zeal reflects better career advisory services that are expected to benefit graduates' early employability, then excluding non-respondents could equally bias university performance. Although non-response rates vary significantly across the institutions in the sample (the average proportion is nearly 10% with standard deviation of 4.8%), we find that the correlation between the percentage of 'unknowns' and the proportion of unemployed and inactive graduates is very low (4.6%). Moreover, when we retained non-respondents in the sample and modelled non-response as an additional category in the micro equation using a multinomial logit regression, we found that the correlation between university rankings calculated with and without these individuals was 99%. This joint evidence suggests that ignoring non-response is not expected to alter significantly the estimated indicators of performance.

[15] For Scottish and Irish students we considered a conversion score of 12 points in their best 5 Higher exams. This would be equivalent to 3 'Bs' and 2 'As'.

minimum of 4% to maximum of 77%). The overall proportion of male students (MALE) is 54%. However, in some institutions the gender split is much more uneven than in others, ranging from 38% to 73%. Significant dispersion is found with respect to the proportion of graduates who took a sandwich year (SANDWICH). Finally, regional unemployment rates vary considerably across the country (UN_RES and UN_UNI).

4.5 Results

Table 4.3 shows the estimation results obtained by fitting Equations (4.1) and (4.3) for the micro and macro models, respectively.[16] The equations include the same adjustment factors, and the *macro* variables X_j are directly derived from their *micro* counterparts, X_i, by taking averages over individuals.[17] The results suggest that *macro* estimates are generally larger and more imprecisely estimated than their *micro* counterpart.[18] Surprisingly, subject studied has a significant impact only in the *micro* regression.[19] We found significant effects, both at the individual and university level, with respect to A-level performance, sandwich courses, prior qualifications, social class, and regional unemployment based on

[16] Graduates' age and degree performance were dropped because highly insignificant at the macro level. As Appendix 4A indicates, this may be due to the high collinearity between A-level score and degree class on one hand, and between mature status and non-A-level prior qualifications, on the other.

[17] Aggregation-by-averaging is the 'implicit' method used to create the university-level information widely used both in related academic research (Johnes and Taylor, 1990) and in reports on university league tables regularly published in the press. Later in the chapter we will show how this approach represents a polar case of a range of alternative aggregation procedures.

[18] These results are in line with previous empirical evidence, including literature on union wage differentials (Booth, 1996), income elasticities of housing demand (Smith and Campbell, 1978), and long-run income elasticities of employment (Lee, Pesaran and Pierce, 1990), where estimates obtained using aggregated data are typically larger than those obtained from individual data.

the student residence prior to university. However, some striking differences emerge with respect to the direction of the estimated *macro* and *micro* effects. For instance, if the university proportion of working class graduates rises by 10 percentage points above its mean level, the predicted proportion of unemployed drops by nearly 5 percentage points. On the contrary, coming from a working class family increases the individual's OLFU probability by 1.2 percentage points.[20] Similarly, the *macro* model predicts that universities with higher proportions of students coming from high unemployment regions have significantly lower proportions of graduates unemployed or inactive, while the micro model predicts the opposite. Aggregation bias provides a possible explanation for these clashing results. In the next section we turn to comparing university rankings.

Table 4.4 shows that the correlation coefficient between the university rankings based on the *macro* indicators $\hat{r}_j$ and the rankings based on the *micro* indicators ϖ_j is 0.63. Graphical comparisons are shown in Figure 4.1. Each point in the diagram represents an institution. In the case of perfect correlation, the points would lie on the 45° line. The diagram shows that only five of the top ten universities hold their positions (area in left-bottom corner delimited by dashed lines). At the opposite end (top-right corner of the diagram), the corresponding ratio is eight in ten. Around the middle of the distribution we find evidence of large movements for some universities. For example, the university with a *macro*

[19] Although not directly comparable, this evidence is at odds with the results found by Johnes and Taylor (1990). The high standard errors may reflect collinearity between the regressors.

[20] For comparability, the micro marginal effects of the control variables were all computed as deviations from the mean (as it is done for continuous variables), rather than as discrete changes in the OLFU probability between 0 and 1.

ranking of 9th moves to a *micro* ranking of 42nd, while the institution ranked 35th according to the macro-based indicators becomes the 3rd when the micro-based indicators are used.

Looking back at the main question that this chapter aims to address, the results discussed above lead us to conclude that university performance indicators and the derived league tables are sensitive to whether university-level data rather than student-level information is used in the analysis. Given that universities are understandably sensitive to their position in the published tables, aggregation issues should at least be acknowledged when employment-related performance indicators are based on university-level information. However, pending further analysis, this conclusion remains tentative. The predictive power of the regressions presented in Table 4.3 is low especially in the micro regression, where the pseudo R^2 does not exceed 2.5%. This is reflected in the high correlation coefficient (0.94) between the 'unadjusted' and the micro-based 'adjusted' rankings (Table 4.4). 'Like-for-like' comparability was bought at the expense of an overly restricted model specification at the micro level. The risk of assessing the effect of data aggregation on the basis of models with low explanatory power is that the impact of an aggregation bias can be overstated. Hanushek, Rivkin, and Taylor (1996) showed that aggregation typically exacerbates the bias due to the omission of relevant explanatory variables. This means that $\hat{r}_j$ and ϖ_j may not be affected in the same way in case of omission of relevant regressors, even if there is a one-to-one correspondence between X_i and X_j. Therefore, the conclusion that league tables are sensitive to data aggregation remains tentative insofar as other sources of bias are not ruled out. To address this issue, one strategy is to search for specifications that maximise the explanatory power of each model, so that $\hat{r}_j$ and ϖ_j can be regarded with more confidence as 'residual' university effects. This implies more latitude and efficiency in the use of our micro level information, once the constraints imposed by like-for-like comparisons are removed. This analysis will be presented in the next section. A second strategy to assess the impact of data aggregation on both parameter estimates and university performance indicators is to use experimental data especially designed to control for any estimation inconsistency other than aggregation. The experiment is carried out by means of a Monte Carlo simulation and will be presented in Section 4.7.

4.6 Comparing *macro* and *micro* outcomes using all the information available

In this section, we estimate more refined specifications of the macro and the micro equations making use of all the information available from the USR. Technical considerations to one side, this seems the logical way to proceed. If one has access to a richer source of information, it is sensible to use it. Clearly, the wealth of information contained in the USR individualised data was only partially exploited in the previous analysis. Not only can we enrich our model with a new set of 'adjusting' factors, but we can also use more efficiently the information presented in Section 4.4. Our objective is to maximise the explanatory power of the regressions. By so doing, performance indicators are expected to gauge more precisely 'residual' university effects on OLFU probabilities.

Table 4.5 shows coefficients, standard errors and marginal effects obtained using 'unrestricted' specifications of Equations (4.1) and (4.3). As in Johnes and Taylor (1990), a measure of the unemployment rate in the district or county where the university is located (UN_UNI) was added to the macro equation to account for unexplained regional labour market effects on graduates first destinations. In the micro model, along with a finer definition of the existing control factors, we additionally included residence dummies and A-level count. The explanatory power of both equations has improved, particularly in the micro regression where the pseudo R^2 has nearly doubled compared to the value obtained in the 'like-for-like' estimation. Although the micro and macro effects of Table 4.5 are generally not directly comparable, contrasting results still emerge when a direct correspondence between the micro and macro outcomes can be found. For instance, the marginal effects associated with A-level Mathematics have opposite signs. The macro marginal effect suggests that if the proportion of leavers with A-level Mathematics increased by 10 percentage points, the proportion of graduates still unemployed or inactive six months after gaining their degree is expected to increase by 1.2 percentage points. Conversely, in the micro regression, having taken maths makes a student 1 percentage point less likely to be in OLFU.

Based on our 'best-fit' regressions, we construct new university rankings, following the same steps as in Section 4.5. Table 4.5 shows that the correlation coefficient between the 'best-fit' macro and micro rankings has only marginally improved (0.64) vis-à-vis the rank correlation obtained when like-for-like specifications were used.

An important question is how confident one can be in the micro and macro rankings. If the dispersion around the average point estimates is high, it may not be possible to 'separate' the performance of different institutions with a reasonable level of confidence. To address the issue, we constructed 90% confidence intervals for the 'best-fit' macro and micro indicators. A horizontal line was drawn through the point estimate of the 'median' university. If the line crosses the confidence bands of institutions occupying lower or higher rank positions than the 'median', then the performance of the former cannot be separated from the performance of the latter. Figure 4.2 shows that the performance of those universities ranked 8th through 41st on the basis of the micro-based indicators ϖ_j cannot be separated at the 90% confidence level. Only the performance of the top seven universities can be 'confidently' regarded as higher relative to the bottom 15 institutions. Estimates are more precise when the macro indicators $\hat{r}_j$ are used.[21] This is equivalent to using the macro residuals $\bar{\hat{r}}_j$ estimated from Equation (4.3).[22] Figure 4.3 shows that only 13 out of 55 institutions cannot be separated.

The persistence of significant discrepancies between micro and macro outcomes, both in terms of parameter estimates and performance indicators, may suggest more convincingly that an aggregation bias exists and it is sizeable. However, this conclusion remains tentative insofar as the impact of other potential sources of

---

21 The 90% confidence interval for $\hat{r}_j$ is calculated as $\hat{r}_j \pm 1.645\sqrt{\sigma_j^2 / N}$ where N is the number of universities, and σ^2_j is the variance of $\hat{r}_j$. The variance σ^2_j is equal to $\sigma^2(1-h_j)$, where h_j is the j^{th} element on the diagonal of the of 'hat' matrix defined as $X(X'X)^{-1}X'$, and σ^2 is the variance of the error term.

22 Rankings are based on P_j, rather than on its logit transformation $log(P_j/(1-P_j)$ which being a monotonic transformation leaves the rankings unaffected.

bias from model mis-specification cannot be completely ruled out. Although an analytic solution to the problem of measuring the aggregation bias originating from the *RA* approach would be more general, its implementation appears exceedingly difficult, owing to the daunting task of linking complex non-linear probability distributions to simpler models based on summary statistics. Empirical solutions, though less general, have the advantage of being much more viable and can suggest, if not final answers, at least overall rules (Hellerstein, 1995). In the next section, we attempt to find more convincing evidence of an aggregation bias by use of Monte Carlo simulation techniques.

4.7 A Monte Carlo simulation

The comparative analysis of the previous sections have highlighted that employment-related league tables based on aggregated data look significantly different from tables obtained from micro data. Although an aggregation bias is a strong candidate to explain the discrepancies observed, we were not able to exclude altogether other possible sources of bias. Monte Carlo simulation techniques represent an attractive empirical solution to neutralise the impact of other sources of mis-specification on the *micro* and *macro* regressions. The idea behind the Monte Carlo experiment presented in this section is to simulate the true data generating process (DGP) of first destination decisions, and then to assess the performance of our *micro* and *macro* models, both in terms of parameter estimates and performance indicators, relative to the 'true' model. The latter is obtained by assuming *full knowledge* of the structural form of the equations, the numerical values of their parameters, and the actual values of the population characteristics.[23]

Given the high number of iterations involved, for practicality reasons we restrict our sample to 18,592 individuals from 51 different universities.[24] For the same reasons, the set of adjustment factors was restricted to a few key variables, including A-level score, gender, social class, method of study (sandwich), and unemployment rate in the region of prior residence. Aggregate information was again obtained by averaging student-level data over individuals for each institution, in a *RA* fashion. Each simulation proceeds as follows:[25]

i. For predetermined values of β (labelled as $\hat{\beta}^{true}$), γ_j ($\hat{\gamma}_j^{true}$) and X_i,[26] we randomly generate individual OLFU probabilities p_i, as follows:

$$p_i = \begin{cases} 1 & if \quad p_i^* > 0 \\ 0 & \quad\quad p_i^* \leq 0 \end{cases}$$

and

$$p_i^* = X_i'\beta + Z_i'\gamma_j + \varepsilon_i \qquad\qquad i = 1, \ldots\ldots, N$$

where p_i^* is a latent random variable measuring the propensity of the individual i towards being unemployed or inactive, X_i is the vector of control variables, Z_i is

23 This Monte Carlo simulation assumes that the 'true' model is the micro model. This seems a plausible standpoint as far as first destination decisions are concerned. However, in general, the superiority of the micro model is not so obvious and should be put to the test (for an example, see Hellerstein, 1995).

24 We drew 30% random samples of graduates from each institution. Consequently, due to the reduced cell size, four small institutions were excluded from the analysis, bringing down the number of universities considered in the simulation to fifty-one.

25 Simulations were carried out with Shazam 7.0 econometric software.

26 The values of β and γ are based on empirical estimates. The values X_i are the actual values of the characteristics of the individuals randomly drawn from each university for the Monte Carlo experiment.

a vector of university dummies and ε_i is a logistic error term generated as $\varepsilon_i = \ln\left(\frac{u_i}{1-u_i}\right)$, with *u~Uniform(0,1)*;

ii. Individual OLFU probabilities p_i are averaged over graduates from each university to obtain aggregate OLFU proportions P_j;

iii. $\hat{\beta}_{micro}$ and $\hat{\gamma}_j$ are estimated from Equation (4.1);

iv. $\hat{\beta}_{macro}$ and the residuals $\hat{r}_j$ are estimated from Equation (4.3);

v. Steps (a) through (d) are repeated for 500 replications.

We then calculated average values (over the 500 replications) of $\hat{\beta}_{micro}, \hat{\beta}_{macro}, \hat{\gamma}_j$ and $\hat{r}_j$, labelled respectively as $\bar{\hat{\beta}}_{micro}$, $\bar{\hat{\beta}}_{macro}$, $\bar{\hat{\gamma}}_j$, and $\bar{\hat{r}}_j$. We also estimated average and peak (absolute) differences between rankings as an additional indicator of the potential alteration produced by data aggregation to the position occupied by single institutions in the tables.

In order to assess the impact of data aggregation, we compare the *macro* and *micro* outcomes of the simulation with the corresponding *true* values predetermined at the outset. Table 4.6 compares the coefficients $\bar{\hat{\beta}}_{micro}$ and $\bar{\hat{\beta}}_{macro}$ (and the derived marginal effects) with the values $\hat{\beta}^{true}$. We note that $\hat{\beta}^{true}$ and $\bar{\hat{\beta}}_{micro}$ are almost identical. This evidence confirms the asymptotic consistency of the maximum likelihood estimator, that is $\bar{\hat{\beta}}_{micro}$ converges to $\hat{\beta}^{true}$ as the number of replication increases. We also observe that the *macro* coefficients clearly overestimate their *micro* counterparts, even though this is not always the case when the marginal effects are considered. Finally, it is clear that the differences between $\bar{\hat{\beta}}_{micro}$ and $\bar{\hat{\beta}}_{macro}$ are not confined to the magnitude, but involve also the direction of the effects. These results seem to confirm the existence of a significant aggregation bias.

With respect to employment-related university performance, Table 4.4 shows the correlation between the rankings based on $\bar{\hat{\gamma}}_j$ and $\bar{\hat{r}}_j$ with the 'true' ranking based on the values $\hat{\gamma}_j^{true}$. Interestingly, while the correlation coefficient between $\hat{\gamma}_j^{true}$ and the *micro*-based indicators $\bar{\hat{\gamma}}_j$ is close to unity (0.998),[27] the correlation between $\hat{\gamma}_j^{true}$ and the *macro*-based performance indicators $\bar{\hat{r}}_j$ does not exceed 0.55. We also find that the position occupied by individual universities in the tables changes, on average, by eleven places when the macro-based indicators $\bar{\hat{r}}_j$ are used instead of the benchmark $\hat{\gamma}_j^{true}$, with a peak variation of thirty-six places. Compared to the correlation coefficients obtained from the 'like-for-like' and 'best-fit' analyses, differences between macro and micro-based rankings have clearly widened. This means that the rank correlation found in the previous sections was inflated (rather than depressed) by the presence of distortions of other nature than data aggregation. Therefore, we conclude that an aggregation bias, probably originating from the inadequacy of the *macro* equation to capture the non-linearity of the underlying *micro* relations, exists and can be sizeable.

---

[27] Given the asymptotic consistency of the ML estimator, the correlation between the 'true' and the *micro* rankings tends to unity as the number of replications increases.

An interesting extension to the analysis presented in this section is to see if the rank correlation between $\bar{\hat{r}}_j$ and the benchmark $\hat{y}_j^{true}$ can be improved upon by using alternative aggregation techniques. Therefore, in the next section we depart from the *RA* approach maintained throughout the chapter and examine two alternative methods of obtaining aggregate predictions from micro data.

4.8 Alternative aggregation procedures

The differences between micro and macro-based outcomes presented so far in the chapter are largely the result of an aggregation bias. The literature on transportation studies reviewed in Section 4.2.1 has shown that predictions obtained from aggregated data heavily depend on the way aggregation is performed. Up to this point, macro-level information was obtained by taking averages of the underlying micro-level data over individuals. It was implicitly assumed that the macro relations are the expression of the behaviour of a fictitious 'representative' individual in the same way as the micro relations describe the decision choices of single agents. However, the *RA* approach is a good approximation only in the unlikely event that the macro unit of observation is the result of the sum of fairly homogeneous micro units. In our case this would mean, for instance, that graduates from the same university are similar in their social background, degree performance, or subject studied. As one departs from this extreme scenario, aggregated behaviour becomes an increasingly poor reflection of actual individual choices.

Given these considerations, an intuitive solution to reduce the aggregation bias introduced by the *RA* approach is to make some allowance at the macro level for the underlying individual heterogeneity. In this section we will analyse in turn two new aggregation methods, which have been used extensively in transportation studies.[28] The first, discussed in Section 4.8.1, uses Taylor series expansions to include second moments (variances) of the explanatory variables in the macro equation (Talvitie, 1973). The second, presented in Section 4.8.2, consists of partitioning the student population into sub-samples on the basis of some key individual or institutional factors with the aim of reducing individual heterogeneity within each segment (Train, 1986).

To assess the effectiveness of these two approaches, we use the same Monte Carlo experiment presented in Section 4.7. The purpose of this exercise is to test whether the correlation between the rankings based on the new indicators and the benchmark $\hat{\gamma}_j^{true}$ exceeds 55%.

4.8.1 Procedures based on Taylor series expansions

A first potential improvement over the *RA* approach is to use a second-order Taylor series expansion of the logistic function (Equation 4.3) about the mean of its explanatory variables. More formally,

[28] See Koppelman (1976) for a survey.

$$P_j^T = P_j + \left[X_j'\beta - \overline{X}_j'\beta\right]\frac{dP_j}{d(X_j'\beta)}\Big|_{\overline{X}_j'\beta} + \frac{1}{2}\left[X_j'\beta - \overline{X}_j'\beta\right]^2 \frac{d^2 P_j}{d\left(X_j'\beta\right)^2}\Big|_{\overline{X}_j'\beta} + R_{2j}$$

(4.4)

where $P_j = L(\overline{X}_j\beta) = \left[1 + e^{\left(-\overline{X}_j'\beta\right)}\right]^{-1}$ is the observed proportion of unemployed or inactive graduates from university j and R_{2j} is the remainder term.[29] Taking expectations on both sides, Equation (4.4) becomes:

$$E\left[P_j^T\right] = \hat{P}_j^T = \hat{P}_j + \hat{\sigma}_j^2 \hat{P}_j\left[1 - \hat{P}_j\right]\left[0.5 - \hat{P}_j\right] + R_{2j} \tag{4.5}$$

where $\hat{\sigma}_j^2$ is given by the product $\hat{\beta}'\Sigma_j\hat{\beta}$ and Σ_j is the variance-covariance matrix of X_i for the university j (Talvitie, 1973).[30] We note that the aggregated OLFU probabilities $\hat{P}_j$ predicted by Equation (4.3) are a special case of $\hat{P}_j^T$ obtained using Taylor-expansion techniques. More precisely, $\hat{P}_j$ equals $\hat{P}_j^T$ when the variances σ_j^2 are equal to zero, that is when all graduates from the same university have endowments of productivity-related characteristics X_i equal to $\overline{X}_i$ (those of the fictitious representative individual). Monte Carlo simulations are run according to the sequence of steps *(a)* through *(e)*) illustrated in Section 4.7. In step (*d*), the new macro indicators $\hat{r}_j^T$ are constructed as the difference between

[29] This method requires knowledge of the second moments (variances and covariances) of the distributions of the explanatory variables. For reasons of practicality related to the high instability of higher order moments, the series are usually terminated after the second term.
[30] In Talvitie (1973) σ^2 is independent of j.

the observed university proportion of OLFU graduates P_j and the proportion $\hat{P}_j^T$ calculated from Equation (4.5). The final indicator $\bar{\hat{r}}_j^T$ is obtained by averaging the values $\hat{r}_j^T$ over the 500 replications. Table 4.4 shows the rank correlation, along with average and maximum rank difference between $\bar{\hat{r}}_j^T$ and $\hat{\gamma}_j^{true}$. Disappointingly, not only has correlation not increased, but its value lies below the threshold of 55% that we were trying to improve.[31] The failure of procedures based on Taylor-series expansions to reduce the aggregation bias is probably due to the fact that variances alone fail to capture adequately the individual heterogeneity in the micro data.

4.8.2 Procedures of classification

Another aggregation method frequently used when the explanatory variables are few and/or take only few values is known as 'classification'. The population is partitioned into classes identified by the values of some key adjusting factors used in the model. In our case, we divide the sample of 18,592 graduates into 4 classes obtained by interacting gender with degree performance.[32] They are: i) males with 'good' degrees, ii) males with 'poor' degrees, iii) females with 'good' degrees, and iv) females with 'poor' degrees. Within each class, macro variables are obtained by the usual *RA* aggregation and used to predict aggregate OLFU

[31] Cameron (1990, p. 216) shows that the remainder term R_{2j} may have very large expected values. Therefore, there is no guarantee that second-order Taylor series approximations are superior to first-order Taylor series expansion (*RA* approximation).

[32] In choosing the classifiers we attempted to maximise class size in order to avoid running into difficulties when the random generating process described in step a) of the Monte Carlo simulation identifies subsets of individuals close to 0% or 100% success rates.

probabilities according to Equation (4.3). Overall OLFU probabilities are then computed as weighted sums of the predicted proportions of unemployed or inactive graduates in each class, with weights equal to the proportion of individuals in each class (Koppelman, 1975a; Train, 1986). As long as individual heterogeneity at the class level is sufficiently lower than at the population level, segmentation can be an effective way of reducing the aggregation bias. More formally,

$$\hat{P}_{jc} = \left[1 + e^{\left(-\overline{X}_{jc}'\hat{\beta}\right)}\right]^{-1} \tag{4.6}$$

where $\hat{P}_{jc}$ is the predicted proportion of unemployed or inactive individuals who graduated from the university j and belong to class c of the population. The overall predicted OLFU proportion is then calculated as,

$$\hat{P}_j^C = \sum_{c=1}^{n} w_{jc}\hat{P}_{jc} \tag{4.7}$$

where n is the number of segments and w_{js} is the proportion of individuals in class c who graduated from university j (Train 1986, p. 101). Performance indicators are calculated by taking the average difference - labelled as $\bar{\hat{r}}_j^C$ - between the values P_j and $\hat{P}_j^C$ predicted by Equation (4.7) over the 500 replications of the Monte Carlo experiment.

Table 4.4 shows that the rank correlation between $\hat{\bar{r}}_j^C$ and $\hat{\gamma}_j^{true}$ is 0.76, that is more than 0.20 higher than the coefficient found when $\hat{\bar{r}}_j$ was used instead. This is reflected in much smaller alterations in the position occupied by individual universities in the rankings.[33] We believe this result provides more convincing evidence that when the aggregation bias is reduced, the gap between macro and micro rankings narrows accordingly.

4.9 Conclusions

This chapter is a first attempt to analyse the effects of data aggregation on the construction of employment-related performance indicators for higher education institutions. The idea was suggested by evidence that most of the published league tables ranking universities according to their success to produce employable graduates are based on university-level first destination information. However, first destination choices are made by individual students according to their aspirations, tastes and ability. Therefore, it seems important that such variety be considered when aggregate measures of university performance are estimated. Unfortunately, when student-level information is unavailable, accounting for individual heterogeneity is generally impossible. In these cases, a convenient but simplistic assumption is that macro relations describe the behaviour of a fictitious 'representative' individual in the same way as the micro relations reflect the

---

33 The underlying assumption is that the rank correlation between $\hat{\bar{r}}_j^C$ and $\hat{\gamma}_j^{true}$ increases monotonically in the number of classes c.

decision choices of single agents. This is, for instance, the assumption underlying most of the studies on performance indicators based on university-level data.

The results presented in the chapter clearly reveal that an *RA* approach to aggregation leads to employment-related university rankings which are significantly different from the rankings obtained from disaggregated data. We argued that these observed differences result from ignoring individual heterogeneity when performance indicators are directly estimated from macro level information. In other words, the observed divergences between macro and micro-based rankings are the result of an aggregation bias. We tested the validity of this conclusion by trying to control for other forms of bias, primarily mis-specification. We believe the evidence produced is sufficient to suggest caution about the validity of league tables based on university averages only. In the absence of individual-level data, the segmentation of the sample into classes could potentially be an effective way to obtain more reliable indicators.

Tables and Figures

Table 4.1 Variables definition (*)

OLFU	Unemployed or out of labour force
SCIENCE	Degree in one of the following subjects: Biology, Other Life Sciences, Physics, Other Physical Sciences, Mathematical Sciences
SOSCI	Degree in Economics, Sociology, Politics, and Other Social Sciences
VOC	Degree in Law, Allied Medicine, Business Studies, Education, Computer Science, and Engineering
SCLOW	Skilled manual, partly skilled, or unskilled parents
ART	Degree in Classics, Modern European Languages and Humanities
ABB+	26 (12) score points or higher in 3 (5) best A-levels (Highers) passes
ALMATH	A-level (Higher) Mathematics
NOALEV	Non-A-level (Higher) entry qualification (BTEC, HND, GNVQ)
MALE	Gender: male
PTIME	Part-time degree
SANDWICH	Sandwich degree
MATURE	Age=21 or older at date of enrolment
INDEP	Independent school
UN_RES	Average unemployment rate (unemployed/economically active) in region/district of prior residence
UN_UNI	Average unemployment rate (unemployed/economically active) in region/district where university is located

(*) Variables are categorical in the micro equation except for UN_RES and UN_UNI.

Table 4.2 Descriptive statistics

variable	mean	standard error			min	max	N
		overall	*between*	*within*			
OLFU	0.164	0.371	0.035	0.369	0.076	0.294	63515
SCIENCE	0.274	0.446	0.096	0.439	0.012	0.569	63515
SOSCI	0.152	0.359	0.105	0.352	0.000	0.731	63515
ARTS	0.222	0.416	0.127	0.404	0.000	0.737	63515
VOC	0.353	0.478	0.174	0.453	0.039	0.771	63515
GOODEG	0.593	0.491	0.064	0.486	0.475	0.774	63515
SCLOW	0.170	0.375	0.041	0.373	0.077	0.243	63515
ABB+	0.307	0.461	0.183	0.412	0.020	0.884	63515
NOALEV	0.113	0.316	0.078	0.304	0.029	0.507	63515
MALE	0.539	0.498	0.074	0.495	0.387	0.731	63515
PTIME	0.014	0.116	0.025	0.112	0.000	0.149	63515
SANDWICH	0.068	0.252	0.230	0.174	0.000	0.968	63515
MATURE	0.111	0.314	0.057	0.308	0.023	0.346	63515
INDEP	0.235	0.424	0.105	0.407	0.007	0.503	63515
UN_RES	0.093	0.039	0.016	0.035	0.076	0.160	63515
UN_UNI	0.110	0.047	0.042	0.000	0.049	0.211	63515
non-resp	0.097	0.296	0.048	0.292	0.019	0.255	70334

Table 4.3 Estimation results: 'like-for-like' regressions

variable	micro		macro	
	β	marginal effect	β	marginal effect
MALE	0.384***	0.051	0.949	0.129
	(0.023)		(0.756)	
SCIENCE	0.164***	0.022	-0.751	-0.102
	(0.026)		(0.506)	
SOSCI	0.215***	0.029	-0.721	-0.098
	(0.031)		(0.557)	
ABB+	-0.297***	-0.039	-1.096***	-0.149
	(0.028)		(0.344)	
ALMATH	-0.198***	-0.026	0.698*	0.095
	(0.025)		(0.379)	
NOALEV	-0.105***	-0.014	-1.943***	-0.264
	(0.038)		(0.690)	
SCLOW	0.074***	0.010	-3.665	-0.498
	(0.029)		(2.265)	
INDEP	-0.071***	-0.009	-0.952	-0.129
	(0.027)		(0.867)	
PTIME	-0.518***	-0.069	5.327*	0.724
	(0.120)		(2.980)	
SANDWICH	-0.135**	-0.018	-0.862***	-0.117
	(0.066)		(0.234)	
UN_RES	1.501***	0.200	-9.134**	-1.242
	(0.305)		(4.466)	
constant	-1.105***		0.022	
	(0.068)		(0.686)	
N	63515		55	
LL	-27766			
R^2			0.46	
adjusted R^2			0.32	
pseudo R^2	0.02			

Table 4.4 Rank correlation coefficients

rank pairs	coefficient		
unadjusted/macro like-for-like	0.599		
unadjusted/micro like-for-like	0.943		
unadjusted/macro best-fit	0.578		
unadjusted/micro best-fit	0.901		
micro/macro (like-for-like)	0.628		
micro/macro (best-fit)	0.642		
Monte Carlo		*avedif*	*maxdif*
unadjusted/true	0.953	3.13	17
true/micro	0.997	0.47	3
true/macro	0.546	10.79	36
micro/macro	0.547	10.78	36
micro/macro (Taylor series)	0.525	11.1	37
micro/macro (classification)	0.763	7.49	28

Figure 4.1 *Macro* versus *micro* university rankings

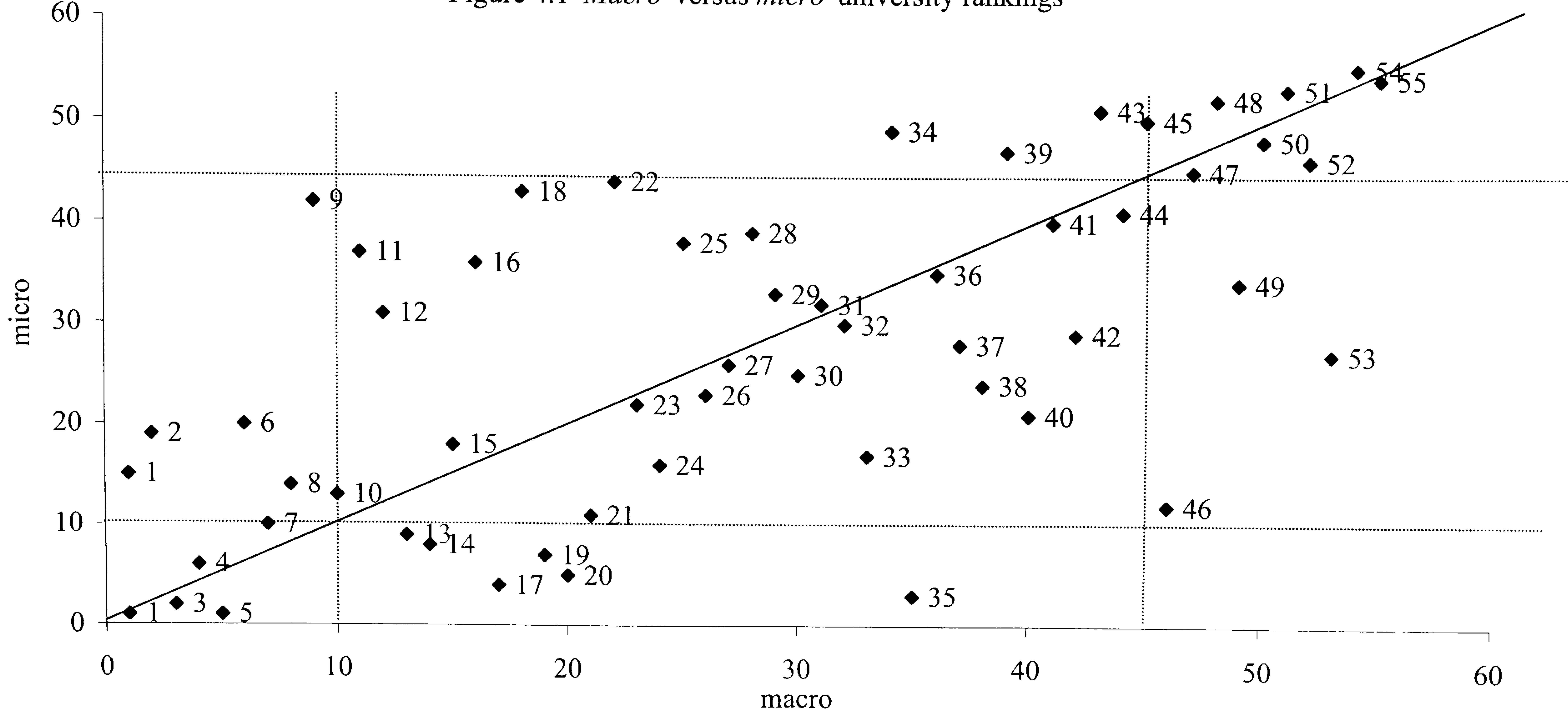


Table 4.5 Estimation results: 'best-fit' regressions

micro				macro			
variable	β		Marginal effect	variable	β		Marginal effect
MALE	0.353	***	0.045	MALE	1.266		0.171
	(0.024)				(0.846)		
CHEM	-0.375	***	-0.048	SCIENCE	-1.208		-0.163
	(0.077)				(0.725)		
LAW	-0.779	***	-0.099	SOSCI	-1.093		-0.148
	(0.08)				(0.697)		
ENGIN	-0.278	***	-0.035	VOC	-0.549		-0.074
	(0.06)				(0.571)		
DFIRST	-0.653	***	-0.083	ABB+	-1.208	***	-0.163
	(0.051)				(0.375)		
ALMATH	-0.075	***	-0.009	ALMATH	0.855	**	0.116
	(0.03)				(0.419)		
NOALEV	-0.11	*	-0.014	NOALEV	-1.776	**	-0.24
	(0.06)				(0.686)		
SC IV	0.142	***	0.018	SCLOW	-4.519	*	-0.611
	(0.044)				(2.39)		
INDEP	-0.134	***	-0.017	INDEP	-1.124		-0.152
	(0.031)				(0.897)		
PTIME	-0.741	***	-0.094	PTIME	5.587	*	0.755
	(0.127)				(2.961)		
SANDWICH	-0.07		-0.009	SANDWICH	-0.726	***	-0.098
	(0.067)				(0.253)		
UN_RES	1.809	***	0.23	UN_RES	-11.628	**	-1.572
	(0.354)				(5.028)		
constant	-1.392	***		UN_UNI	1.463		0.198
	(0.12)				0.983		
				constant	(0.45)		
N	63515					55	
LL	-27058						
R^2						0.4897	
adjusted R^2						0.328	
pseudo R^2			0.047				

Figure 4.2 90% confidence intervals for the micro-based performance indicators

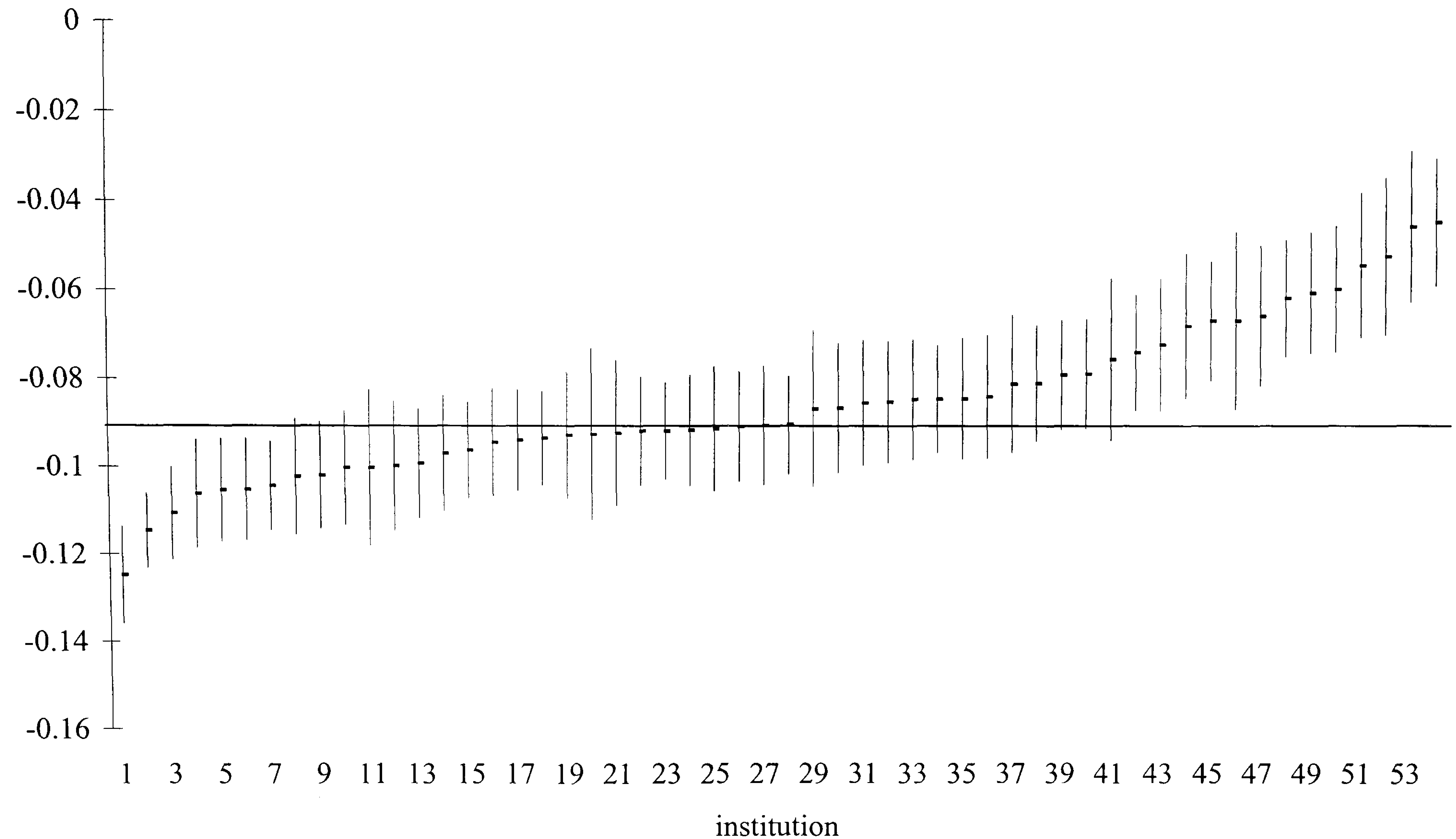


Figure 4.3 90% confidence intervals for the macro-based performance indicators

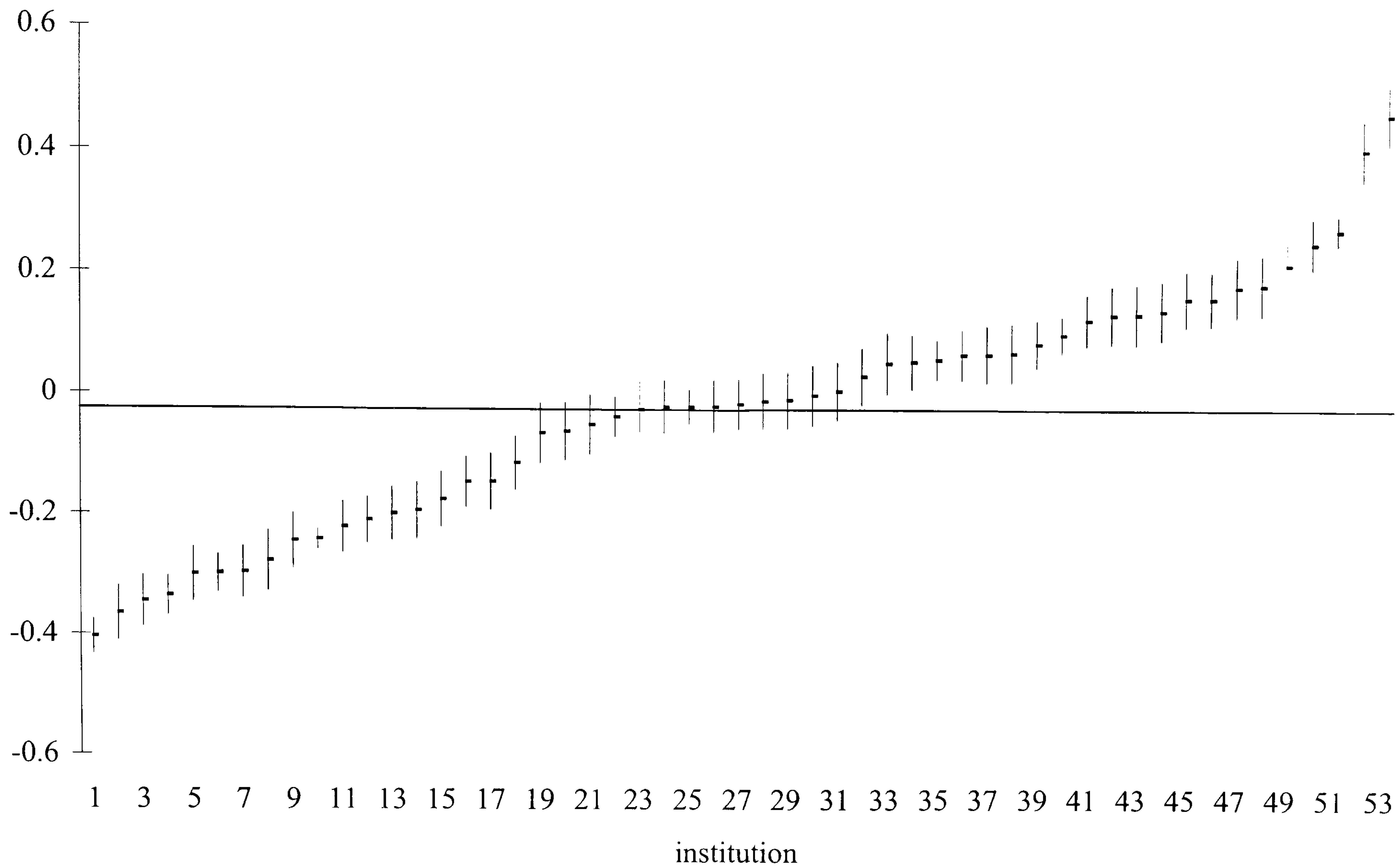


Table 4.6 Monte Carlo estimates

variable		micro		macro	
	$\hat{\beta}^{true}$	$\bar{\hat{\beta}}_{micro}$	me	$\bar{\hat{\beta}}_{macro}$	me
MALE	0.367 *** (0.041)	0.365	0.145	0.576	0.068
ABB+	-0.344 *** (0.050)	-0.345	-0.137	-1.125	-0.133
SANDWICH	-0.069 (0.117)	-0.068	-0.027	-0.496	-0.059
ALMATH	-0.136 *** (0.042)	-0.138	-0.054	0.735	0.087
SCLOW	0.021 (0.053)	0.019	-0.007	-3.571	-0.424
UN_RES	1.855 *** (0.554)	1.888	0.751	-4.499	-0.534
constant	-0.993 *** (0.119)	-1.002		-1.076	

Appendix 4A Correlation matrix

	OLFU	MALE	MATURE	SCIENCE	SOSCI	ARTS	VOC	GOODEG	ABB+	NOALEV	ALMATH	INDEP	SCLOW	PTIME	SANDWICH	UN_UNI
OLFU																
SEX	-0.02															
MATURE	-0.04	-0.35														
SCIENCE	0.08	0.12	-0.39													
SOSCI	0.00	-0.15	0.16	-0.39												
ARTS	0.21	-0.63	-0.07	0.00	0.07											
VOC	-0.19	0.49	0.17	-0.32	-0.44	-0.77										
GOODEG	-0.10	0.12	-0.43	0.08	0.20	0.08	-0.23									
ABB+	-0.05	0.37	-0.63	0.23	0.01	-0.01	-0.12	0.78								
NOALEV	-0.23	-0.07	0.76	-0.22	-0.23	-0.29	0.47	-0.40	-0.41							
ALMATH	-0.06	0.74	-0.48	0.30	-0.29	-0.57	0.42	0.22	0.53	-0.08						
INDEP	0.14	0.03	-0.53	0.25	0.10	0.29	-0.41	0.68	0.69	-0.50	0.19					
SCLOW	-0.16	0.15	0.24	-0.15	-0.05	-0.26	0.30	-0.68	-0.63	0.23	-0.03	-0.85				
PTIME	-0.21	-0.22	0.58	-0.23	-0.01	-0.14	0.24	-0.23	-0.21	0.66	-0.24	-0.41	0.08			
SANDWICH	-0.40	0.31	-0.03	-0.23	-0.19	-0.45	0.57	-0.01	-0.11	0.15	0.21	-0.26	0.29	0.10		
UN_UNI	-0.06	0.04	0.16	-0.12	-0.12	-0.19	0.28	-0.38	-0.12	0.31	0.01	-0.44	0.34	0.41	-0.09	
UN_RES	-0.27	-0.11	0.45	-0.21	-0.11	-0.20	0.33	-0.36	-0.22	0.61	-0.10	-0.55	0.31	0.78	-0.01	0.72

Chapter 5

Differences in the occupational earnings of UK graduates by degree subject: evidence from the USR, 1980-1993

5.1 Introduction

To this point, the Thesis has focused extensively on the determinants of graduates' early careers, and particularly on graduates' employability upon graduation. A second dimension of labour market success that has attracted considerable attention among labour economists are the pecuniary rewards of a first degree qualification. The empirical literature has shown that there are substantial positive returns to an undergraduate university degree in the UK. Blundell *et al.* (2000), for instance, using National Child Development Survey (NCDS) data find that the average return to an undergraduate degree, in terms of wages, with respect to individuals aged 33 with two or more A-level passes who did not continue into higher education, was 17% for men and 37% for women in 1991. However, the majority of the Mincer-type earnings regressions found in the return-to-education literature estimate the average rate of return to a university degree and do not control for field of study. In a recent review of the literature Chevalier *et al.* (2002) show that the private rate of return to a university degree is likely to differ substantially by degree subject. Moreover, to the best of our knowledge, we are not aware of any UK studies that also attempt to model directly student self-selection into university subjects, despite acknowledging the importance of the potential endogeneity of subject choice.[1]

This chapter aims to contribute to the empirical return-to-education literature by estimating occupational earnings *premia* by subject studied using alternative methods to control for the potential endogeneity of subject choice. First, we use

[1] Blundell *et al.* (2000) acknowledge the issue and use a 'matching and proxying' method to reduce the impact of selectivity. However, this methodology relies on rather restrictive assumptions, namely that the wealth of information considered in the analysis is sufficient to control for selectivity biases.

OLS estimation techniques widely used in related research for the UK. This standard approach is not only interesting for comparison purposes with previous research, but represents also a useful benchmark to assess the existence and size of a potential selection bias. Second, this first set of results is contrasted with estimates obtained from propensity score matching methods which have become an increasingly popular technique in the evaluation literature (Rosenbaum and Rubin, 1983). Although substantially different, both methods hinge on the assumption that selection is driven solely by observable factors. Finally, we introduce a third approach consisting of a simultaneous model of graduate earnings and subject choice (Lee, 1983), which also allows for self-selection through unobservable attributes.

This concept of heterogeneous returns across degree courses is particularly relevant over time. As more individuals experience higher education, just holding a university degree becomes a weaker distinguishing mark for students and a less informative screening device for the talent available to employers, if not supplemented by information on the graduates' awarding university, field of study, or degree class obtained. On the grounds that the economic return to a degree depends on the demand and supply for that specific university specialisation, our multi-cohort analysis over the period 1980-1993 is also expected to provide useful information on the trends of the graduate labour market in the UK.

The outline of the chapter is as follows. In Section 5.2, we report the findings of previous studies on the differences in graduate earnings by degree subject in the UK. Section 5.3 illustrates three alternative modelling strategies. Section 5.4

discusses some issues regarding occupational earnings data, while Section 5.5 describes the salient features of the sample. Section 5.6 presents the three sets of results for male graduates obtained from OLS, propensity score matching, and the simultaneous equation model, respectively. Section 5.7 presents the results for female graduates, and examines some of the potential causes of gender differences in subject *premia.* Finally, Section 5.8 concludes summarising the main results.

5.2 Previous literature

In comparison to the rich literature investigating the return to education in the UK (see, for instance, Harmon and Walker 1995, 1999, 2000 and Blundell *et al.,* 2000), there are only few studies which analyse differences in graduate earnings by degree course.

Dolton and Makepeace (1990) using data from the Survey on 1980 Graduates and Diplomates conducted by the Department for Employment in the UK find that the average earnings of Commerce graduates (including Accounting, Business & Management, Economics and Law) were higher than the earnings of graduates in other disciplines.

Belfield *et al.* (1997) use survey data on the 1985 and 1990 graduate cohorts to investigate differences in 1996 average salaries by subject of degree. The authors find that the relative ranking of degree subjects based on average male salaries remained unchanged for the two cohorts, with Social Sciences ranked first followed in order by Science, Humanities and Education.

Chevalier (2000) uses 1996 current salaries and pooled survey data on the 1985 and 1990 graduate cohorts to estimate the relative earnings *premia* by degree course considering 12 different subject groups. He finds that graduates from Mathematics and Social Science earned respectively 6% and 2% more than graduates in Education, while Humanities graduates earned 12% less.

Chevalier *et al.* (2002) using more broadly defined subject groups and 1996 earnings data, estimate a similar specification for the 1980 graduate cohort and find earnings differences with respect to Education of +8.4%, +11.6% and +6.8% for Science, Social Science and Humanities, respectively. The corresponding figures for the 1995 cohort using 1999 average salaries were +17.9%, +16.8% and +5.4%, showing an improvement of Science and Social Sciences relative to Education.

Naylor *et al.* (2002) use USR individualised data on the 1993 graduate cohort and find significant differences in inter-occupational earnings across degree subjects. The most economically rewarding subjects were Law, Computer Science, and Allied Medicine while the less rewarding were Agriculture, Humanities, and Classics & Literature.

Lissenburgh and Bryson (1996) use data from the 4th wave of the Youth Cohort Study 3 (YCS) and find that graduates from Science, Mathematics and Engineering earned 9% more than other graduates.

Blundell *et al.* (2000) using UK data from the National Child Development Study (NCDS) find that graduates in Economics, Accountancy and Law performed significantly better in terms of hourly wages (at age 33) than those with undergraduate degrees in other subjects such as Arts, Chemistry & Biology and Other, the residual category.

Harkness and Machin (1999) use data from the General Household Survey (GHS) between 1980 and 1995 focusing on the return to degree subject for full-time workers. They consider four broad subject groups (Arts, Science, Social Science and Other) and find, *inter alia*, that the return to a degree in Arts for males was about 10% lower than in the other fields.

Blackaby, Murphy, and O'Leary (1999) using UK data from the 1993-1995 Labour Force Surveys (LFS) find that male graduates' earnings vary significantly across degree subjects. In particular, after controlling for a number of personal, job and demographic characteristics, the authors find that graduates from Economics, Accountancy, Law and Management did better than their peers in other subjects, especially compared to Other Social Sciences and Arts.

Walker and Zhu (2001) using LFS data for the period 1993-1999 find that '[..there are no systematic trends in returns by subject nor is there any tendency for them to converge]' (p. 37). The study shows marked differences in the return to an undergraduate degree across subjects. Graduates in Economics, Law, 'Health' (i.e. medical related) and Mathematics ranked at the top of the earnings scale while graduates in Arts performed substantially worse (with a negative mark-up with respect to students with at least two A-level passes who did not continue in higher education).

5.3 Methodology

This section describes in turn three alternative approaches used to estimate graduate occupational earnings *premia*[2] by subject of degree: i) the 'proxying and matching' method (OLS), ii) the 'propensity score matching-average treatment on the treated' model (PSM-ATT), and iii) a simultaneous equation model of earnings determination and subject choice (MNL-OLS).

5.3.1 Selection on observable factors: the 'proxying and matching' method (OLS)

A common method to ascertain earnings differences by degree subject is to estimate by OLS an earnings function specified as:[3]

$$y_{ij} = \sum_{j=1}^{J} S_{ij}\theta_j + X_i'\beta + \varepsilon_i \tag{5.1}$$

where y_{ij} is the natural logarithm of the earnings of individual i who studied subject j, X_i is a vector of individual attributes which may affect both subject choice and occupational earnings, S_{ij} is a dummy variable which takes value one if the individual i graduated in the subject j and zero otherwise, and θ_j is the earnings *premium* of graduating from subject j relative to the default case. As observed by Blundell *et al.* (2000), this is tantamount to matching individuals on the basis of the index $X_i'\beta$ and to assuming equality of θ_j's across individuals.

---

[2] In this chapter, by 'earnings *premia*' we refer to 'log-earnings' *premia*.

[3] This is the method used in all the studies reviewed in Section 5.2.

The OLS model does not require any distributional assumption on ε_i, but it does require orthogonality between ε_i and X_i.

5.3.2 Selection on observable factors: the PSM-ATT model

An alternative method to estimate subject *premia* is to compare occupational earnings for individuals who graduated in one subject with 'matched' individuals who studied for a different degree course. This framework considers the subject of study as the treatment that the individual receives and aims to assess the causal effect of this treatment on the outcome variable, namely occupational earnings. The direct comparison between individuals in different treatment groups may be misleading because they may differ systematically in their observable and unobservable characteristics. In their seminal paper, Rosenbaum and Rubin (1983) suggested the use of propensity scores matching procedures to solve the issue of sorting due to observable factors. The propensity score is defined as the conditional probability of receiving the treatment given an individual's characteristics:

$$p(X_i) \equiv \Pr\{S_{ij} = 1 \mid X_i\} \qquad (5.2)$$

where S_{ij} are dummy variables which take value one if individuals graduated in the subject j and zero otherwise, and X_i is the vector of conditioning factors that we observe. In our case, the propensity scores are computed by estimating binary logit models of subject choice for each of the broad course categories defined in Section 5.4 using Economics and Business graduates as the reference group.[4] Under the assumption that differences between individuals affecting the outcome are entirely captured by their observed characteristics X_i,[5] the average treatment effect on the treated (ATT) can be estimated as follows:

$$ATT = E\{y_{i1} \mid P(X_i), S_{ij} = 1\} - E\{y_{i0} \mid P(X_i), S_{ij} = 1\} \tag{5.3}$$

where y_{i1} and y_{i0} are the occupational earnings of graduates in subjects 1 and 0, respectively. In words, individuals with the same value of the propensity score $P(X_i)$ but different treatments S_{ij} , act as controls for each other and the average difference between their earnings equals the ATT.

Compared to the *proxying and matching* method (OLS), the assumption of equality of the subject *premia* θ_j's across individuals is relaxed. In fact, here the earnings *premia* are computed as the average of the earnings differences between 'matched' pairs of treated-untreated individuals. This method is non-parametric and does not require any distributional assumption on the unobservables.

---

4 This implies the breakdown of our sample into *J-1* sub-samples of graduates in each year. Each sub-sample include the 'treated' individuals, that is the individuals who graduated in a specific subject and the 'untreated' individuals, i.e. the individuals who graduated in the reference subject (Economics and Business) and, therefore, received a different 'treatment'. Strictly speaking, we are evaluating the differential impact of alternative treatments.

5 This is known as the Conditional Independence Assumption (CIA), formally: $y_{0i} \perp S_{ij} \mid P(X_i)$ The other necessary assumption is the so-called 'common support' assumption: '[All treated agents have a counterpart on the non-treated population and anyone constitutes a possible participant: $0 < P(S_{ij} = 1 \mid X_i) < 1$]' (Blundell and Costa Dias, 2002, p. 22).

5.3.3 Selection on observable and unobservable factors: a simultaneous equations model of earnings determination and subject choice

The two methods illustrated in Sections 5.3.1 and 5.3.2 rely on the assumption that treated and untreated individuals differ only with respect to observable attributes. Hence, these approaches neglect the possibility of *self-selection* with respect to unobservable characteristics. If individuals choose the degree subject by maximising their individual utility, the students enrolled in the different subjects are those who have comparative advantages, i.e. those for whom the choice turns out to be optimal. If this sorting effect is not fully accounted for by observable attributes, both the OLS and the PSM-ATT estimates of the earnings *premia* by subject are likely to be biased.

The econometric framework we use to address the self-selection on unobservables in a polychotomous choice model was developed by Lee (1983). Below, we report the main features of the model.

Let us assume that the utility of the student i in the subject j (V_{ij}), with j=1,...J, depends on individual's characteristics (Z_i) and on a idiosyncratic unobservable term u_{ij}, reflecting for instance individual's preferences over degree subjects, in the following way:

$$V_{ij} = Z_i \delta_j + u_{ij} \qquad j=1,...,J. \tag{5.4}$$

Occupational earnings are generated according to the following process:

$$y_{ij} = \sum_{j=1}^{J} S_{ij}\theta_j + X_i\beta + \varepsilon_i \tag{5.5}$$

where y_{ij} are the log-earnings of individual i who read subject j, X_i is a vector of individual attributes and S_{ij} is a dummy variable which takes value one if the individual i graduated in the subject j and zero otherwise, and θ_j's are the subjects earnings *premia*, which are the primary focus of our analysis. The *selection bias* arises from the correlation (ρ_j) between the stochastic components u_{ij}'s and ε_i, that is, between the unobserved individual's characteristics affecting subject choice and those affecting occupational earnings. If the model does not account for this correlation, the subject dummies may simply pick up the effect of the individual's unobserved characteristics rather than the 'true' earnings *premium* associated with the subject studied. For instance, because the type of occupation is typically correlated with the subject studied at university, individuals with a preference for certain jobs will be more likely to choose those subjects more related to their preferred occupation. After graduation, the individual will be more likely to be observed in his/her preferred occupation, which in turn will affect his/her earnings. Therefore, the earnings *premia* by subject studied may also capture the effect of idiosyncratic occupational preferences.[6] Another possible source of selection bias could be the higher preference for some non-pecuniary

---

[6] Arcidiacono (2002), for instance, shows that ability sorting observed across majors is mainly determined by idiosyncratic preferences for certain jobs and certain subjects rather than by expected performance or expected earnings.

characteristics of the job by graduates in certain subjects. This may explain, for instance, why graduates in Humanities are at the bottom of the earnings scale.[7]

Lee (1983) suggests that the model above can be estimated using maximum likelihood methods, under some specific assumptions on the distributions of the stochastic terms u_{ij}'s and ε_i. Here we assume that the u_{ij}'s are independent and identically Gumbel-distributed, while $\varepsilon_i \sim N(0,\sigma^2_\varepsilon)$. The form of the log-likelihood function in this specific case is shown in the Appendix 5B. The attractiveness of estimating simultaneously a Multinomial Logit-OLS model (MNL-OLS hereafter) is that the model does not impose restrictions on the correlations between the unobservables affecting the subject choice and the individual's earnings, which are jointly estimated along with the other parameters of the model.[8]

In brief, we estimate simultaneously a MNL model for the subject choice and an earnings regression in which the degree subject appears as one of the explanatory variables. In this way, we aim to estimate the differences in graduate earnings by degree subject corrected for *self-selection bias.*

Model identification

Even when the vectors Z_i and X_i coincide, the different functional forms of Equations (5.4) and (5.5) (non-linear *vs* linear) are sufficient to identify the simultaneous equations model. To ensure that identification does not rely exclusively on the specific functional forms adopted, we seek to find variables that affect subject choice but do not influence earnings (i.e. we look for an

[7] Daymont and Andrisani (1984), for instance, find that students in Humanities have weaker preferences for pecuniary job characteristics (earnings).

‘economic’ identification). We restricted our focus on A-level curriculum and age. From a theoretical point of view, the type of secondary school curriculum is a pre-requisite (in terms of type of pre-university knowledge or entry requirements) for some university courses, and should therefore affect subject choice (for instance, see Altonji, 1993). Van de Werfhorst *et al.* (2002) provide some empirical evidence for the UK. By contrast, we do not have a strong *a priori* about its importance for graduates’ earnings, once controlled for the degree subject.[9] As for the student’s age, Becker’s (1993) human capital theory predicts that younger individuals, who have a longer expected working life, have higher returns to education and also to more selective and lucrative subjects requiring higher effort. Davies and Guppy (1997), for instance, found that older students were less likely to enter more lucrative fields. Therefore, we considered age as a potential candidate for the identification of our simultaneous equation model. Although we acknowledge that age may affect earnings through work experience accumulated prior to university and, therefore, may be a weak identifying variable, we expect this to be less true when occupational earnings are used compared to actual salaries.

Once potential ‘candidates’ are identified, we test whether the chosen variables actually affect subject choice but not earnings determination. OLS earnings regressions are run separately by year of graduation. In any year, the set of identifying variables (ID.Vs henceforth) were selected from the set of ‘candidates’ by excluding the least significant (the one with the highest p-value above the

---

8 Therefore, we do not impose the independence of u_{ij}*'s* and ε_i, which would allow the separate estimation of the earnings equation and the subject choice models.

9 A-level Mathematics is an exception and, therefore, is not included among the potential identifying variables.

threshold of 0.10), and performing a likelihood ratio (LR hereafter) test for the validity of the restriction. If the test is passed, the second least significant variable (with p-value greater than 0.10) was excluded and a new cumulative LR test performed. This procedure was reiterated until the cumulative LR test for the joint omission was rejected. We investigated the sensitivity of the estimated *premia* to alternative identification strategies (functional form *vs* functional form and ID.Vs). The results are reported in Appendix 5C.

5.4 Earnings data and control variables

The analysis presented in this chapter is based on USR data on individual university students who graduated from 'pre-1992' UK universities between 1980 and 1993. Unfortunately, the USR does not include information on employed graduates' individual salaries. However, the First Destination Record does contain detailed information on the (self-reported) type of occupation held by graduates six months after graduation.[10] As in Naylor *et al.* (2002), we were able to match the individual's reported occupation to the corresponding (gender-specific) 3-digit SOC of the NES.[11] Occupational earnings were then computed as the average

[10] In the First Destination Record of the USR occupations are classified into more than 120 categories contained into 6 'major' groups and 28 'minor' groups.

[11] The New Earnings Survey is an annual survey of pay and hours of work, by far the largest of its kind in the UK, producing 2 million observations between 1975 and 1998 (Bell and Elias, 2000). Unlike most earnings surveys, the information is collected from employers rather than from employees. This is generally accepted as producing more accurate estimates of earnings, since employers are perhaps less inclined to misrepresent employees' earnings than employees themselves.

gross weekly pay of individuals employed full-time (in the same occupation) in the year following graduation.[12]

The length of time since it was first conducted in 1975 makes the NES an ideal source to study long-term trends in pay. However, the change in the NES occupational classifications that took place in 1990 has required a significant coding effort to ensure consistency/continuity over time.[13] Prior to 1990, the coding scheme used in the NES was the Key List of Occupations for Statistical Purposes (KOS), which consisted of 404 occupations arranged into 18 main groups. From 1990 onwards occupational data were coded to the Standard Occupational Classification (SOC), consisting of 371 unit groups contained into 77 'minor', 22 'sub-major', and 9 'major' groups. To help bridge these two coding schemes, individuals were classified under both schemes in 1990. However, the match was fairly imprecise. Some KOS occupations were scattered across a number of SOC occupations and vice versa. This, in turn, meant that for some occupations there was a jump in the average earnings series in 1990 due to this reclassification process (Bell and Elias, 2000).

Notwithstanding the reclassification, the existence of a dual coding in 1990 and the fact that the USR classification remained substantially unchanged over the period 1980-1993, enabled us to achieve a satisfactory level of consistency in the earnings series before and after 1990. However, the quality of the match was significantly higher for males. For female graduates, the repercussions of the 1990 change in the NES coding schemes on the coherence of the SOC-to-USR mapping

[12] The age range considered in the computation of average earnings is 18-63 for men and 18-59 for women.

of occupations over time, and ultimately on the calculation of average earnings, have been more serious. This was largely due to the generally smaller sample size of occupational groups for females both in the NES and in the USR, which caused average earnings to be more volatile over time.[14] To some extent, changes in occupational segregation by gender and in women's participation over time may have contributed to exacerbate the discontinuities in the female occupational earnings series. For these reasons, in the next sections the use of female earnings will be restricted to the period 1990-1993.

The use of occupational earnings has advantages and disadvantages compared to individual starting salaries. A clear disadvantage is the loss of any intra-occupational variation in pay. On the other hand, occupational earnings have the advantage of being a better proxy for career earnings and, therefore, a better measure of the lifetime rate of return to a university degree, compared to starting salaries. We are aware that occupational earnings based on graduates' first destination information collected six months after graduation may be only weakly correlated to later career earnings. However, as already pointed out in Chapter 2 (Section 2.3), there is evidence suggesting that this is not necessarily the case (Dolton and Makepeace, 1992; Purcell and Pitcher, 1996; McKnight, 1999).

---

[13] The SOC-to-USR and the KOS-to-SOC mappings and the calculation of average occupational earnings were kindly provided by Abigail McKnight. We are grateful to her for making the data available to us.

[14] A smoothing of the break in the female earnings in 1990 would have required aggregating occupational groups up to a level which did not guarantee enough inter-occupation variation to estimate our earnings equations efficiently.

We focus on five broadly defined subject areas:[15]

1. *Science* (including Life Sciences, Physical Sciences and Mathematical Sciences);
2. *Hi-Tech* (including Computer Science, Engineering and Technology);
3. *Eco-Bus* (Business and Economics);
4. *HSS* (Humanities and Other Social Sciences);
5. *Other* (a residual and rather heterogeneous category).

The purpose of this chapter is to estimate relative earnings *premia* by subject studied over time. Therefore, it is crucial to control for a number of individual factors, which are also expected to affect graduate earnings. They are:

(i) *Family background.* As in Chapter 2, graduates were grouped into six social classes: SC I , SC II, SC IIINM, SC IIIM, SC IV-V and SC OTH;

(ii) *Schooling background.* This includes controls for A-level grades, number of A-level passes by broad subject field, curriculum breadth, and the type of school attended;

(iii) *Personal characteristics.* These include age (mature student status), marital status, and residence prior to university.

---

[15] See Appendix 5A for a detailed definition of the broad subject categories and the A-level groups. The remaining variables used in the analysis are defined as shown in Appendix 2A. Due to the complexity of the model and the number of parameters to be estimated we were not able to consider a finer definition of academic subjects. A similar level of aggregation is used both in the articles reviewed in Section 5.2, and in international studies on college major's choice correcting for sample selection (for instance, see Berger (1988) and Rochat and Demeulemeester (2001).

Our dependent variable is the natural logarithm of average nominal gross weekly occupational earnings. In the next section we present the main features of the sample.

5.5 Sample and summary statistics

Given the focus of the chapter, we consider only those graduates who reported an occupation six months after graduation.[16] Amongst the employed, we apply more stringent selection criteria than those already discussed in Chapter 3: besides non-respondents, non-UK students, medical students, and individuals from non-traditional (non A-level) entry routes to higher education, we further excluded part-time graduates.[17] After selection, cohort size ranges between 17,100 and 21,300 for males and between 16,200 and 18,900 for female graduates.

Figures 5.1 and 5.2 show the pattern of average occupational earnings by broad subject over time for male and female graduates, respectively.[18] A few interesting points emerge. First, male earnings are considerably more stable. Even at this high level of aggregation, there is a clear break in the 1990 earnings series for female graduates in Eco-Bus and Hi-Tech courses. This evidence substantiates the choice to use female earnings from the post-reclassification period only. Second, female average earnings are significantly lower. Given the relevance of earnings

[16] Like all the studies reviewed in Section 5.2, in this chapter we do not address the issue of the potential biases due to self-selection into employment and survey non-response.

[17] Part-time students are excluded from the analysis since they have previous work experience and probably different early career outcomes relative to full-time students. In any year, part-timers generally represent less than 2% of the sample.

differentials between males and females at the policy level, in Section 5.7.1 we will take a closer look at the gender earnings gap. Third, the relative ranking of subjects looks very different between men and women. Figure 5.1 shows a sizeable earnings gap between Eco-Bus and the remaining subjects (excluding Other), which has increased over time. On the contrary, Figure 5.2 suggests that Eco-Bus female graduates have generally lower average earnings than any other groups.

It is also interesting to look at the spread of occupational earnings in each subject over time. Table 5.1 shows that during the entire sample period, Hi-Tech graduates (both male and female) were those with the lowest variance in occupational earnings, generally followed by Science graduates. This may be simply due to the fact that Hi-Tech graduates find employment in occupations with similar pay levels. However, it could also be the case that Hi-Tech degrees are more 'specialist', in the sense that graduates with this specialisation are observed in a relatively narrow range of occupations. Consequently, there might be little variation in occupational earnings across individuals. As a crude test of the latter hypothesis, Tables 5.2 (males) and 5.3 (females) show for each subject the distribution of graduates across broad occupation and sector categories. We find that the ranking of subjects by degree of specialisation is identical for males and females for these broad types of occupation. For instance, according to this criterion, Eco-Bus subjects are the most 'specialist' because 40.8% (30.9%) of male (female) graduates go into Accounting occupations. The corresponding proportions in the modal occupations for Science, Hi-Tech, and HSS graduates

[18] Females' occupational earnings pre-dating the change in the NES coding system are plotted

are 10.2% (14.9%), 26.3% (21.6%), and 12.5% (18.5%), respectively. These results are in line with the findings reported in Dolton and Makepeace (1990).[19]

With respect to the sector of employment, the ranking of subjects varies considerably by gender. Amongst males, Hi-Tech graduates are the most 'specialist', with the Engineering & Construction (EC) sector attracting over 50% of the employed. In the other subject groups, the modal sector attracts 15%, 37% and 21% of Science, Eco-Bus, and HSS graduates, respectively. This ranking order is nearly reversed for female graduates: HSS ranks first with 44.2% of their graduates employed in the Public Administration, Health and Education (PHE) sectors, followed by Science (43.9% in PHE), Hi-Tech (38.0% in EC) and Eco-Bus (30.2 % in Accounting).

The evidence presented in Tables 5.2 and 5.3 seems to confirm the intuition that the degree of specialisation of university subjects is negatively correlated with the dispersion of occupational earnings, especially among males.

Figures 5.3 (males) and 5.4 (females) show the proportion of graduates by broad degree subject and year. For males, the proportion of graduates in Hi-Tech degrees increased steadily in the first half of the 1980s, and remained rather stable until 1990, when numbers started to decline. The proportion of HSS graduates fell in the early 1980s but has since increased, especially in the 1990s. After a decline in the period 1980-1982, the proportion of Science graduates has generally increased during the 1980s. In 1989, numbers started to fall but during the 1990s

---

using dashed lines.

[19] Dolton and Makepeace (1990) construct an 'entropy' score measuring the degree of specialisation of academic subjects in terms of first-job destinations. The most 'specialist' degrees were found to be, in decreasing order, Education, Law, Health, Engineering, Economics and Accounting.

trends remained stable. Finally, the proportion of Eco-Bus graduates had slightly increased throughout the sample period.

Figure 5.4 shows that concentrations of female graduates across university subjects were remarkably stable over time. Moreover, it confirms the existence of significant gender differences in the concentration across subjects, which are well documented in the literature and have been already commented extensively in previous chapters. HSS degrees are by far the most popular courses among women followed by Science, while the least popular are Hi-Tech degrees.

The next section presents and compares the estimation results from the three alternative models discussed in Section 5.3. The much shorter earnings series for female graduates (1990-1993 only) leads us to focus primarily on the estimation on males. The estimation results for females will be presented separately in Section 5.7 along with an in-depth analysis of the main gender differences.

5.6 Results for males

Table 5.4 reports, for each cohort, three sets of subject *premia* (with standard errors) estimated from Equations (5.1), (5.2-5.3), and (A.5.1) in Appendix 5B, respectively. To aid interpretation, and to help understand the dynamics of relative earnings *premia*, the results are also shown graphically in Figures 5.5 (OLS), 5.6 (PSM-ATT), and 5.7 (MNL-OLS). The discussion that follows examines the predictions of each method in turn and is largely based on the graphical analysis.

5.6.1 OLS

A first important result is that in all years the differences in occupational earnings by degree subject are highly statistically significant (Table 5.4, part I).[20] Consequently, the returns to a university education estimated by standard Mincerian earnings regressions, which typically do not control for subject of study, is only an average measure and it fails to capture the marked differences in returns that exist across broadly defined subjects.

Second, Figure 5.5 shows that the relative rank of degree subjects is stable over time, except for Science and Hi-Tech whose ranks swap position in 1980 and 1990.

Third, it is clear that over the whole period Eco-Bus graduates (the reference category) had positive earnings *premia* with respect to graduates in other disciplines.[21] This result is in line with the findings of most of the UK literature reviewed in Section 5.2.[22]

Fourth, it is also evident that the average return to a university degree was much more similar across subjects in 1980. Relative to Eco-Bus, the (negative) *premium* of graduating in Science, Hi-Tech or HSS was -2.4%, -3.1% and -3.4%, respectively. The size of these *premia* widened between 1981 and 1987 when the

[20] Only in 1982, 1983, and 1989 are the relative subject *premia* of 'Other' subjects not statistically significant.

[21] The only exception is represented by the category labeled as 'Other', whose relative *premia* are positive from 1989 onwards. Given the high level of heterogeneity of this group, we do not comment this result.

[22] Our estimates are not directly comparable with the findings of most of the literature reviewed in Section 5.2, since we use a different definition of earnings and only observe university students. However, the qualitative results, and especially the relative ranking of degree subjects, closely replicate the findings of those studies.

gaps reached -8.3%,-8% and -9.7% (for Science, Hi-Tech and HSS).[23] Whilst Science and Hi-Tech gained ground on Eco-Bus between 1987 and 1991, in the last two years the gap widened again. By contrast, the relative rank of graduates in HSS worsened in the first half of the 1980s when the *premium* stabilised at about -10% with respect to Eco-Bus.

5.6.2 PSM-ATT

Table 5.4 (part II) shows the average treatment effects on the treated (with standard errors). The matching procedure is based on the 'single nearest neighbours' method.[24] The relative (to Eco-Bus) earnings *premia* of Science, Hi-Tech and HSS are all statistically significant and negative, in line with the OLS results. Figure 5.6 shows that the subject *premia* for Science and Hi-Tech run parallel to, but are systematically lower than, their OLS counterpart. Therefore, relative to OLS, PSM-ATT yields higher negative earnings *premia* for Science and Hi-Tech graduates. On the contrary, with the exception of 1984 and 1987, the negative earnings *premia* of HSS graduates are lower when PSM-ATT techniques are used. These trends have a bearing on the dynamics of the relative ranking of subjects over time. The most striking result is the change in the relative position of HSS, which is no longer systematically at the bottom of the earnings scale. In some years, HSS graduates have enjoyed (positive) earnings *premia* with respect to both Science and Hi-Tech.

---

23 Here, we are mainly concerned with the description of the trend in earnings mark-ups due to degree subjects and do not have the ambition to explain their causes.

5.6.3 MNL-OLS

The relative subject *premia* estimated by MNL-OLS are shown in Table 5.4 (part III).[25] The results look very different from those obtained from the OLS and PSM-ATT methods. For instance, there is a high and statistically significant correlation between the unobservables u_{ij} affecting subject choice and ε_i influencing occupational earnings for Hi-Tech graduates (Table 5.6 shows that ρ is always greater than 0.74 in absolute value). With the noticeable exception of the 1981, 1988, and 1991-93 cohorts, the correlation is generally positive suggesting that those factors inducing enrolment into Hi-Tech courses also tend to command higher earnings in the labour market. As a consequence, the OLS and the PSM-ATT estimates of the Hi-Tech earnings *premium* are biased upwards ('positive selection bias'). However, as noted above, in 1981, 1988, and 1991-93 the direction of the bias is reversed (i.e. 'negative selection'). Since ρ is the correlation between two sets of unobservables, it is difficult to offer an economic interpretation of the direction and magnitude of the bias.[26] A tentative explanation for the 1981 and 1991-93 peaks is the 'specialist' nature of Hi-Tech courses. In fact, in these two periods the UK manufacturing sector, which employs the bulk of Hi-Tech graduates, suffered from a severe crisis. In 1981, the

---

[24] Estimation is performed using the *psmatch2* Stata command (Leuven and Sianesi, 2000). We use single nearest-neighbor matching estimator without caliper and with replacement imposing a common support. Standard errors are bootstrapped using 500 replications.

[25] The selected ID.Vs by year are reported in Table 5.5, which also shows the results of the LR test for their exclusion from the MNL model (Bound *et al.*, 1995). It is worth noting that the set of ID.Vs is rather stable over time. Furthermore, the rejection of the exclusion restrictions in the MNL model confirms the high significance of the ID.Vs for the subject choice. Finally, the results reported in Appendix 5C show that the estimated earnings *premia* are generally robust to changes in the identifying strategy.

[26] We tried to explain these peaks using business cycle indicators both at an aggregate and sectoral level. However, the results were inconclusive, in the sense that we found no correlation between

sector experienced its worst crisis since the start of the economic recession in the late 1970s. This partly contributed to accelerate the secular expansion of the service sector. Similarly, in the period 1990-92, the engineering industry experienced a 10% contraction.[27] These negative sector trends may have had the effect of magnifying the negative OLS *premia* of Hi-Tech relative to Eco-Bus, because the negative sector effect is wrongly ascribed to the subject studied. Clearly, the other 'outlier' found in 1988 is more difficult to justify under this line of reasoning, since the late 1980s were a period of steady expansion for the manufacturing sector.

A second explanation is linked to evidence produced by Nicholson and Souleles (2001) in a study of physicians' income. They find that income prediction errors might be very different according to the speciality undertaken. These errors may depend on unanticipated market and practice changes. For instance, physicians practising in a market where the demand for their services or the payments from health insurers increased, each earned about $29,000 more than expected. This shows how market factors can change the actual realisations of income, and individuals who expected a substantially lower (higher) income might turn out to receive big unexpected gains (losses) from a specific occupation. This might also explain why the sign of ρ changes dramatically year-on-year. Finally, the use of occupational earnings based on first occupation could magnify the size of subject's *premia* and exacerbate their volatility over time. In fact, on one hand, the distribution of new graduates across occupations and jobs is likely to be more

---

the estimated *premia* and the cyclical components of the real GDP and employment obtained using Hodrick-Prescott filtering techniques.

[27] D. Grow, 'Recession in Engineering worse than 1990'. The Guardian, 3 October 2001.

sensitive to economic fluctuations or to sector-specific shocks compared to the whole stock of graduates in the labour market (see Chapter 3, Figure 3.2). On the other hand, occupational earnings do not account for intra-occupational differences in pay by level of experience. This aspect of the data is expected to magnify inter-occupational differences vis-à-vis starting salaries, because the gradient of pay levels to work experience can be very different (e.g. teaching and engineering).

To a certain extent, the time profile of the Science *premium* mimics the pattern of Hi-Tech. The smoother time profile of the earnings *premium* may be explained by the fact that Science graduates generally find employment in both the service and manufacturing sectors.[28] As shown in Chapter 3 (Figure 3.3), the service sector has been less sensitive to economic fluctuations than the manufacturing sector. Consequently, Science graduates were less exposed to the effects of economic fluctuations than Hi-Tech graduates. Unlike OLS, MNL-OLS results suggest that Science graduates earned on average more than Eco-Bus graduates in some years.

We observe a negative *premium* associated to HSS relative to Eco-Bus throughout the sample period except for 1980. However, the magnitude of the *premium* is smaller (the average for the whole period falls from –8.8 to –3.7%). This is due to the existence of a positive selection bias for Eco-Bus courses. It is also interesting to note that HSS graduates are not systematically at the bottom of the earnings distribution, as generally indicated by the OLS results.

Finally, ρ is always positive and generally significant for Eco-Bus degrees. In addition, the size of the bias is particularly high in 1980-81, 1988, and 1991-1993.

---

[28] This is due to the more 'generalist' nature of Science degrees (see Table 5.3).

Again, the concomitant positive selection bias for Eco-Bus and the negative selection bias for Hi-Tech commented above, may explain the widening of the relative earnings *premium* associated with the latter degrees in these periods.

Given the volatility of the MNL-OLS *premia*[29] in Figure 5.7 we no longer observe any systematic ranking of subjects over time.[30] This result is in line with the evidence found by Walker and Zhu (2001).

5.7 Results for female graduates

In this section we present the results for female graduates who left university between 1990 and 1993, using average occupational earnings based on the SOC90 only. Table 5.7 reports the earnings *premia* by subject estimated using the OLS, PSM-ATT, and MNL-OLS model, respectively. The results of the first two models both indicate that Hi-Tech female graduates earn a *premium* over individuals with Eco-Bus degrees, while no significant differences in expected earnings are found for HSS graduates. With respect to Science courses, the earnings *premium* estimated with OLS is no longer significant when the PSM-ATT model is used.

Like for males, when relative subject *premia* are estimated using the MNL-OLS approach, some of the coefficients change dramatically. In this case, the main differences with respect to the OLS and PSM-ATT estimates are found for Hi-Tech in the period 1990-92 when we observe sizeable and negative earnings

[29] In Appendix 5D we show a graphical analysis comparing OLS, PSM-ATT and MNL-OLS earnings *premia* for each subject over time.

premia vis-à-vis Eco-Bus. These results may suggest that women who take Hi-Tech degrees constitute a smaller and more selective segment of the female ability distribution than do males. Furthermore, Hi-Tech female graduates may be strongly motivated individuals, highly committed to succeed in traditionally male careers (where they may be discriminated against). If these factors remain unaccounted for, then the previous estimates of the Hi-Tech *premium* for female graduates are expected to be biased upwards. It is interesting to note the dramatic change of the Hi-Tech coefficient from large and negative to large and positive in 1993. Given that the break coincides with the start of the economic recovery after the depression of the early 1990s, the result may suggest that the selection bias is strongly counter-cyclical. This conclusion seems in line with the idea that the boosting effect of ability and motivation factors on labour market success is higher during periods of economic downturns when demand is slack.[31] In fact, when the economy stagnates for any given level of education, individuals with low ability or motivation will be the first to be fired by a firm. In these circumstances, not only are female Engineers expected to be, on average, better off than their male counterparts, but they are also better off compared to female graduates from other disciplines.

30 In Appendix 5E we performed likelihood ratio tests for the equality of earnings *premia* by subject between consecutive years. The results reject at the 1% statistical level the null hypothesis of equality for every pair of consecutive years.

31 The availability of female earnings data prior to 1990, and particularly in the expansionary period of the late 1980s, would have been useful to test the hypothesis of a link between the bias and the business cycle.

5.7.1 Gender differences

The descriptive analysis discussed in Section 5.5 pointed to pervasive and persistent gender differences both in the distribution of degree courses and in the average occupational earnings by subject. Not only are women's earnings consistently lower than men's, but also the relative ranking of subjects looks significantly different. For instance, while average pay levels of Eco-Bus male graduates exceed those of other subject groups (excluding 'Other'), Eco-Bus female graduates are generally at the bottom of the earnings scale. Table 5.8 shows that the overall earnings gap[32] is in the region of 30% and it is rather stable over time. However, the size of the gap varies between subjects. Eco-Bus female graduates earn around 40% less than male graduates in the same subject, while for Hi-Tech graduates the gap is around 26%.

This evidence is based on simple averages and does not take into account gender differences in human capital and other (observable) productivity-related characteristics. However, even when these differences are controlled for, the size of the gap remains largely unaffected. Using propensity score matching methods,[33] male and female graduates in the same subject group were matched on a large set of observable characteristics.[34] The results, shown in Table 5.9, reveal that the earnings ratio between a female graduate and her 'nearest male neighbour' varies between 60% to 78% across subjects, with relatively little

[32] This is computed as the difference in the log-average occupational earnings of males and females.

[33] To our knowledge, the application of propensity score matching methods to the analysis of the gender wage gap has not been attempted before. Although strictly speaking gender cannot be regarded as a 'treatment', statistical matching techniques represent nonetheless an attractive tool to isolate the effect of gender on occupational earnings. We are grateful to Dr Barbara Sianesi and Prof. Andrea Ichino for useful suggestions.

change over the period 1990-1993. As before, Eco-Bus stands out as the subject group with the larger gap (women earn on average 40% less than men). At the other end, Hi-Tech is the field where the gap is narrower (women earn on average 28% less). These results imply that the gap is almost exclusively due to differences in unobservable factors.[35]

Given that the size of the gender earnings gap varies considerably between subjects, it is not surprising that when the relative earnings *premia* for male graduates (Table 5.4) are compared with the corresponding female coefficients (Table 5.7) we hardly find any similarities. Gender differences are not just limited to the size of the coefficients θ_j, but extend also to their sign and statistical significance. For instance, the expected earnings of HSS graduates lag significantly behind that for Eco-Bus graduates, but only for males. Moreover, the result is substantially unaffected by the type of model used. This evidence may offer an explanation as to why HSS courses largely remain a preserve for females. Hi-Tech *premia* are another example of marked gender differences: OLS and PSM-ATT estimates indicate that Hi-Tech female graduates earn, on average, significantly more than Eco-Bus graduates, while the *premium* is negative for males. When selection on unobservables is taken into account, a sizeable bias is found for both genders. However, the direction of the bias is generally negative for males (OLS massively underestimates the Hi-Tech *premium*), and positive for females (the OLS *premium* is massively inflated).

---

[34] In the set of matching variables we have also included degree class, university type and sector of employment.

[35] To test for the reliability of the results, we performed for each subject 'Oaxaca-Blinder' decompositions of the earnings gap and found that differences in observable characteristics explained between 2% and 6% of the gap.

It is widely agreed that sex-based differences in school content are a key factor to explaining gender differences in earnings. Brown and Corcoran (1997) using US data from the 1984 Survey of Income and Program Participation (SIPP) and from the 1972 National Longitudinal Study (NLS) find that subject differences account for about 20% of the male-female wage gap among college graduates, after controlling for demographic characteristics and work experience. Machin and Puhani (2003) use data from the UK and German Labour Force Surveys of 1996 and find that subject of degree explains between 9 to 19 percent of the overall gender wage gap.

From a policy perspective, this evidence has been used to recommend that gearing more women towards male-dominated subjects could reduce the gender wage gap (Machin and Puhani, 2003). However, as noted in some studies, this policy will have little effect if i) males' and females' wages differ because of underlying differences in innate talent and preferences and not because of differences in academic curricula, or ii) women rationally choose not to specialise in male-dominated fields because they receive lower returns than men from those fields (Brown and Corcoran, 1997; Paglin and Rufolo, 1990).

These arguments highlight the existence of a two-way link between gender differences in educational choices and gender differences in labour market rewards. Therefore, controlling for the endogeneity of subject choice is an important aspect of the analysis. In empirical studies of the male-female wage gap, it is common to ascribe to labour market discrimination the proportion of the gap that is not explained by gender differences in productivity-related characteristics. However, if unobserved gender differences in innate preferences or ability cause educational dissimilarities, neglecting the endogeneity of subject choice would overstate the amount of discrimination by attributing part of labour market differences to demand as opposed to supply factors. On the other hand, to the extent that the perception of entry barriers to certain jobs and/or the 'glass-ceiling' effects[36] influences subject choice, the impact of discrimination will be underestimated (Polachek, 1975).

To shed light on the issues, an interesting test would be to look at how the subject choices of female graduates change in response to changes in the relative subject *premia* and in the labour market return to other productivity-related characteristics. The test is carried out by constraining the coefficients θ_j and β in the MNL-OLS female regression (Appendix 5B, Equation (A.5.1)) to be equal to the corresponding values estimated from the male equation, and by looking for any significant change in the female predicted probabilities of subject's choice. This has important policy implications. If women become more likely to take male-dominated courses, then reducing the wedge between men and women's rewards to identical types of human capital could, through a decline in educational and occupational segregation, have far-reaching effects on abating the wage gap. On the other hand, if educational choices are inelastic to changes in labour market returns, gender differences in subject concentrations are likely to stem from underlying differences in tastes and preferences. Females might at the same time be perfectly rational and decide to choose apparently lower return but also less-demanding subjects/occupations in terms of time commitment and effort. The results (Table 5.10) show that shifts in females' subject choices are unlikely to

36 'Glass-ceiling' means that while men and women may enter occupations on the same terms,

occur. In other words, removing gender differences in the labour market returns to personal and course-related characteristics does not affect female students' educational choices. This implies that men's and women's wages differ because of underlying differences in innate talent and preferences and not because of differences in academic curricula.

5.8 Summary and concluding remarks

This chapter has presented alternative estimates of the occupational earnings of UK university graduates by degree subject using USR and NES data from 1980 to 1993. The analysis is innovative because it does not limit itself to recognize that subject choice may be endogenous to the determination of earnings *premia*, but attempts to correct directly for student self-selection into degree courses. The results obtained from a standard OLS approach are contrasted with estimates from propensity score matching techniques, which correct for selectivity through observable characteristics only, and simultaneous equation models of earnings determination and subject choice, which also account for selectivity through unobservables. We find that, irrespective of the estimation technique used, the differences in occupational earnings by degree subject are in general statistically significant. This confirms that returns to a university degree estimated by standard earnings regressions controlling only for the level of educational attainment, and not for the subject studied, have to be considered only as the average return of a

---

women find it harder to advance trough the ranks (Dolton *et al.*, 1996).

university degree, with marked differences even across broadly defined subjects. Our main findings from the three approaches are:

(a) *OLS*: male graduates in Economics and Business had positive earnings *premia* with respect to graduates in other subjects for the whole period. This result is consistent with the existing UK literature. The relative ranking of subjects has not changed between 1980 and 1993: after Eco-Bus subjects, the second most remunerative subject area is Hi-Tech (including Engineering, Technology, and Computer Science), followed by Science, and Humanities and other Social Sciences. The results look very different for females. Eco-Bus graduates are generally at the bottom of the earnings scale, trailing behind Hi-Tech, Science and HSS graduates;

(b) *PSM-ATT*: using a semi-parametric matching approach to compare the earnings of individuals from different subjects we find that Eco-Bus male graduates rank still first in the earnings scale. However, compared to OLS, the negative *premia* associated with Science and Hi-Tech degrees are higher throughout the sample period. By contrast, HSS courses are no longer ranked last. For female graduates, the most noticeable difference relative to OLS is the fact that the positive Science *premium* is no longer significant;

(c) *MNL-OLS*: when taking into account student selection into field of study based on both observable and unobservable student characteristics, the dynamic profile of the estimated earnings *premia* by degree course becomes considerably more erratic, with the consequence that no stable ranking of subjects emerges over time. Furthermore, we generally observe positive selection for graduates in Economics & Business and Hi-Tech (except for periods of economic downturn for the latter group), while no selection was generally found for HSS graduates. The evidence of nonzero correlation between unobservable factors driving both subject choice and occupational earnings cast doubts on the reliability of estimates based on methods where selectivity runs only through individuals' observable characteristics. In fact, earnings differences due to individual unobserved characteristics may be wrongly ascribed to the subject of degree. Moreover, earnings *premia* are likely to change over time, thus affecting the relative ranking of subjects even between consecutive years. As a consequence, studies focusing on specific cohorts of graduates may give only a very short-term account of the relative economic return to different degree subjects.

Finally, we used the MNL-OLS model to investigate gender differences in occupational earnings, and in particular to test the plausibility of a policy aimed at reducing the gender wage gap (Machin and Puhani, 2003) through gearing more women towards high-pay male-dominated subjects like Hi-Tech courses. Women may rationally choose not to specialise in male-dominated fields because they receive lower returns than men from those fields (Brown and Corcoran, 1997; Paglin and Rufolo, 1990). One way to test this claim is to equalise the relative returns to subject studied and other human capital factors across genders and see if females' subject choices change accordingly. We find that the females' educational orientations are unlikely to shift. Men and women wages differ because of underlying differences in innate talent and preferences and not because of differences in academic curricula.

Tables and Figures

Figure 5.1 Male gross weekly occupational earnings by subject and year

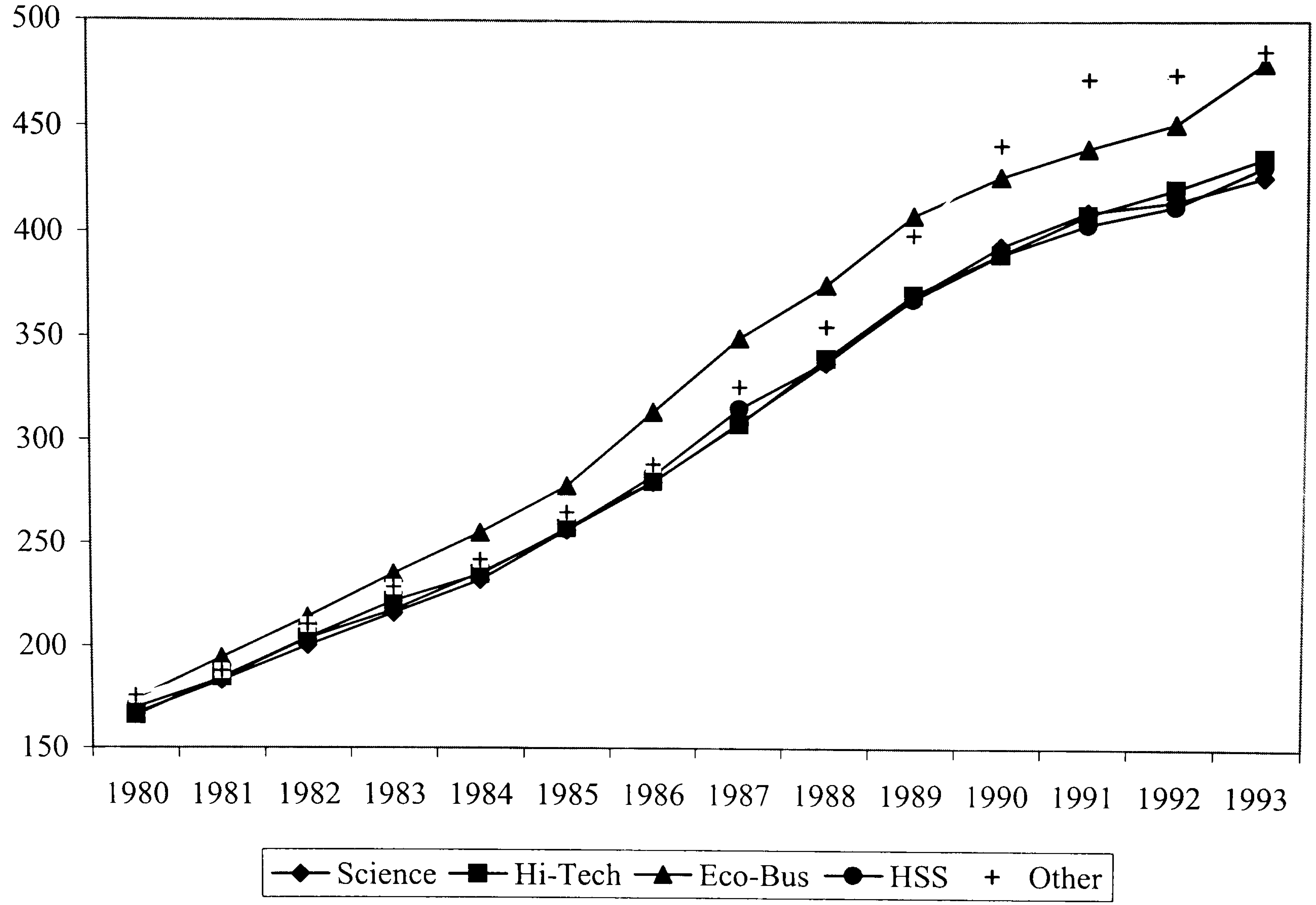


Figure 5.2 Female gross weekly occupational earnings by subject and year

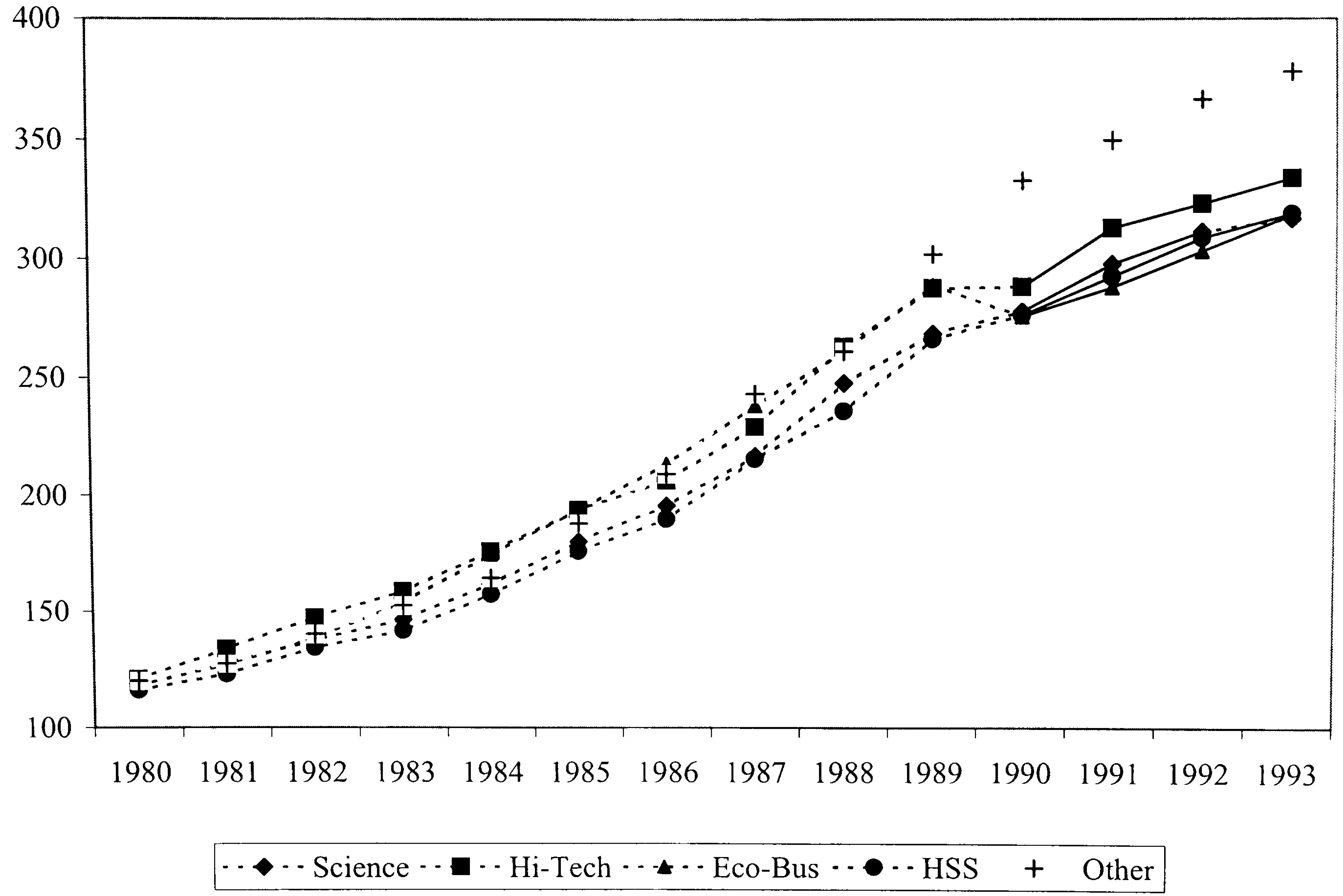


Table 5.1 Mean, standard error and coefficient of variation of nominal gross weekly occupational earnings by gender, subject and year

year		Males						Females					
		Degree subjects					Total	Degree subjects					Total
		Science	Hi-Tech	Eco-Bus	HSS	Other		Science	Hi-Tech	Eco-Bus	HSS	Other	
1980	mean	166.82	165.82	173.89	169.70	175.58	170.19	117.91	121.00	118.07	116.05	120.05	117.75
	s.d.	24.90	14.48	24.30	30.14	27.07	25.26	22.69	18.22	19.79	23.47	24.86	23.49
	c.of v.	0.15	0.09	0.14	0.18	0.15	0.15	0.19	0.15	0.17	0.20	0.21	0.20
	N	5030	3994	2703	4291	4899	20917	3063	262	789	6280	3953	14347
1981	mean	183.05	184.60	194.39	184.04	187.70	186.11	127.32	134.03	127.52	123.13	127.62	125.68
	s.d.	28.58	17.56	28.68	34.49	28.32	28.34	25.40	19.41	24.60	25.68	24.76	25.31
	c.of v.	0.16	0.10	0.15	0.19	0.15	0.15	0.20	0.14	0.19	0.21	0.19	0.20
	N	4655	4040	2647	4215	4558	20115	2962	306	940	6638	3983	14829
1982	mean	200.08	203.90	214.19	203.63	210.37	205.87	137.64	147.47	138.54	134.52	140.25	137.28
	s.d.	34.35	19.8	35.01	40.93	34.85	33.85	28.14	23.73	26.55	29.46	28.21	28.67
	c.of v.	0.17	0.10	0.16	0.20	0.17	0.16	0.20	0.16	0.19	0.22	0.20	0.21
	N	4353	4221	2653	3913	4479	19619	3155	365	1110	6619	4066	15315
1983	mean	215.99	222.11	235.49	217.64	229.01	223.18	146.64	158.83	154.08	141.97	152.87	147.33
	s.d.	41.16	25.11	40.10	47.89	40.95	39.94	30.54	27.46	30.23	31.27	31.38	31.41
	c.of v.	0.19	0.11	0.17	0.22	0.18	0.18	0.21	0.17	0.20	0.22	0.21	0.21
	N	4856	4800	2885	3990	4556	21087	3396	455	1271	6830	4547	16499
1984	mean	232.22	235.04	255.59	235.68	242.38	238.62	162.05	176.04	174.83	157.42	164.64	162.31
	s.d.	41.52	24.31	45.48	52.32	40.62	41.48	33.09	26.67	35.90	37.03	33.87	35.43
	c.of v.	0.18	0.10	0.18	0.22	0.17	0.17	0.20	0.15	0.21	0.24	0.21	0.22
	N	5076	5202	2753	3981	4279	21291	3500	558	1328	6862	4333	16581
1985	mean	256.69	257.66	278.12	256.86	265.47	261.46	180.37	193.84	193.85	176.10	188.02	182.04
	s.d.	44.93	26.61	51.36	59.61	45.76	46.28	39.75	33.87	42.20	43.40	40.76	42.07
	c.of v.	0.18	0.10	0.18	0.23	0.17	0.18	0.22	0.17	0.22	0.25	0.22	0.23
	N	5200	4973	2806	3859	3684	20522	3578	585	1389	7047	4108	16707
1986	mean	280.00	280.49	314.17	283.32	288.73	286.93	195.62	206.20	213.33	189.76	209.13	198.37
	s.d.	59.82	36.79	67.53	71.65	53.18	58.90	43.02	36.28	45.13	45.52	48.48	46.30
	c.of v.	0.21	0.13	0.21	0.25	0.18	0.21	0.22	0.18	0.21	0.24	0.23	0.23
	N	4966	4794	2677	3926	3640	20003	3572	584	1471	6918	3984	16529

Table 5.1 (continued)

		Males						Females					
year		Degree subjects					Total	Degree subjects					Total
		Science	Hi-Tech	Eco-Bus	HSS	Other		Science	Hi-Tech	Eco-Bus	HSS	Other	
1987	mean	308.82	308.05	349.68	315.54	326.26	318.42	216.90	229.41	238.40	215.49	243.36	225.26
	s.d.	67.20	44.00	79.90	83.44	61.21	68.38	48.21	41.44	52.48	53.66	57.21	54.37
	c.of v.	0.22	0.14	0.23	0.26	0.19	0.21	0.22	0.18	0.22	0.25	0.24	0.24
	N	5258	4907	2701	4002	3631	20499	3605	611	1482	6942	4192	16832
1988	mean	337.69	339.88	375.36	338.34	355.26	346.44	247.92	263.35	261.50	235.99	261.52	248.40
	s.d.	72.77	45.30	78.56	85.82	71.13	72.00	50.38	44.59	53.62	55.13	54.14	54.62
	c.of v.	0.22	0.13	0.21	0.25	0.20	0.21	0.20	0.17	0.21	0.23	0.21	0.22
	N	5231	4886	2739	4135	3637	20628	3586	565	1639	6792	4284	16866
1989	mean	369.43	370.64	408.43	368.26	398.96	380.04	268.81	287.92	289.34	266.56	302.25	279.19
	s.d.	79.88	47.39	83.98	89.84	82.14	78.15	60.61	47.21	57.58	63.82	71.84	65.85
	c.of v.	0.22	0.13	0.21	0.24	0.21	0.21	0.23	0.16	0.20	0.24	0.24	0.24
	N	4602	4937	2767	4097	3429	19832	3421	662	1792	6620	4126	16621
1990	mean	393.51	389.78	426.78	389.69	441.63	405.04	278.09	288.38	275.88	276.62	332.88	291.64
	s.d.	89.80	59.14	87.10	100.54	101.76	89.96	60.19	50.21	58.87	67.53	83.00	73.08
	c.of v.	0.23	0.15	0.20	0.26	0.23	0.22	0.22	0.17	0.21	0.24	0.25	0.25
	N	4305	4604	2790	3706	3176	18581	3364	647	1799	6245	4118	16173
1991	mean	409.90	408.48	440.25	403.93	472.92	424.23	297.84	313.03	288.21	292.54	349.53	308.65
	s.d.	92.59	60.99	95.71	110.46	112.19	98.40	70.64	57.95	66.59	78.32	82.73	79.92
	c.of v.	0.23	0.15	0.22	0.27	0.24	0.23	0.24	0.19	0.23	0.27	0.24	0.26
	N	3907	3908	2662	3608	3043	17128	3359	634	1861	6345	4251	16450
1992	mean	414.80	420.72	451.76	412.77	474.90	432.18	311.40	323.14	303.19	308.87	366.59	324.20
	s.d.	100.38	72.04	106.29	115.80	117.87	105.75	82.97	63.61	71.02	87.97	95.19	90.06
	c.of v.	0.24	0.17	0.24	0.28	0.25	0.24	0.27	0.20	0.23	0.28	0.26	0.28
	N	4023	3898	2777	3939	3199	17836	3582	681	1776	6815	4442	17296
1993	mean	426.54	435.29	480.21	431.46	486.07	447.95	316.69	333.81	318.50	318.79	377.90	333.60
	s.d.	111.67	84.48	122.92	124.11	118.15	115.14	88.45	77.39	82.09	92.72	98.48	95.31
	c.of v.	0.26	0.19	0.26	0.29	0.24	0.26	0.28	0.23	0.26	0.29	0.26	0.29
	N	4302	4017	2818	4277	3216	18630	4011	714	1860	7603	4701	18889

Table 5.2 Index of subject's specialisation by occupation and sector: males

	Occupation			Sector		
subject	modal occupation	%	relative rank	modal sector	%	relative rank
Science	ACC	10.2	4	EC	15.1	4
Hi-Tech	EC	26.3	2	EC	50.6	1
Eco-Bus	ACC	40.8	1	ACC	37.4	2
HSS	TP	12.5	3	PHE	21.3	3

Table 5.3 Index of subject's specialisation by occupation and sector: females

	Occupation			Sector		
subject	modal occupation	%	relative rank	modal sector	%	relative rank
Science	ACC	14.9	4	PHE	43.9	2
Hi-Tech	EC	21.6	2	EC	38.0	3
Eco-Bus	ACC	30.9	1	ACC	30.2	4
HSS	TP	18.5	3	PHE	44.2	1

Figure 5.3 Proportion of male graduates by 'broad' degree subject and year (%)

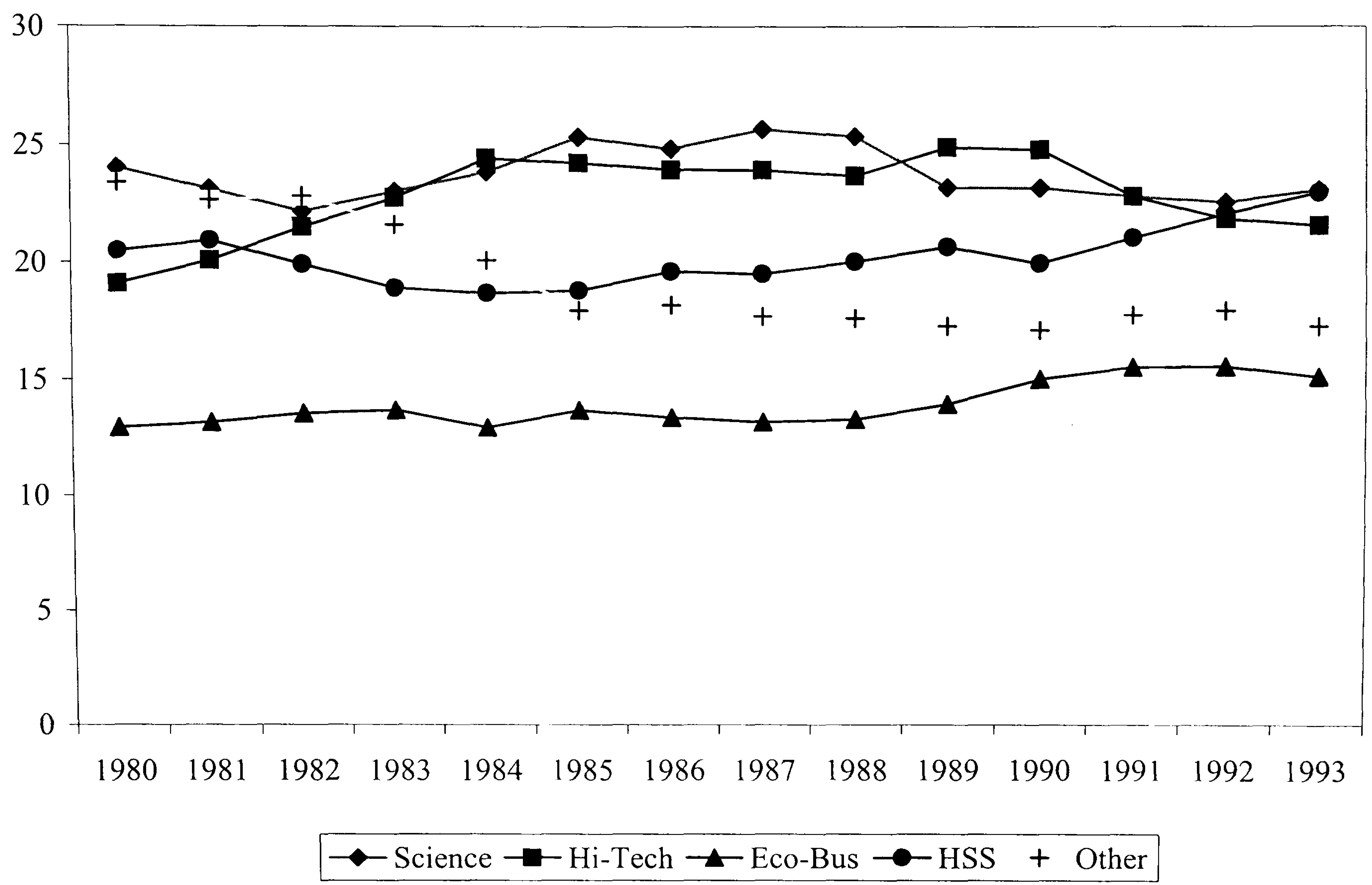


Figure 5.4 Proportion of female graduates by 'broad' degree subject and year (%)

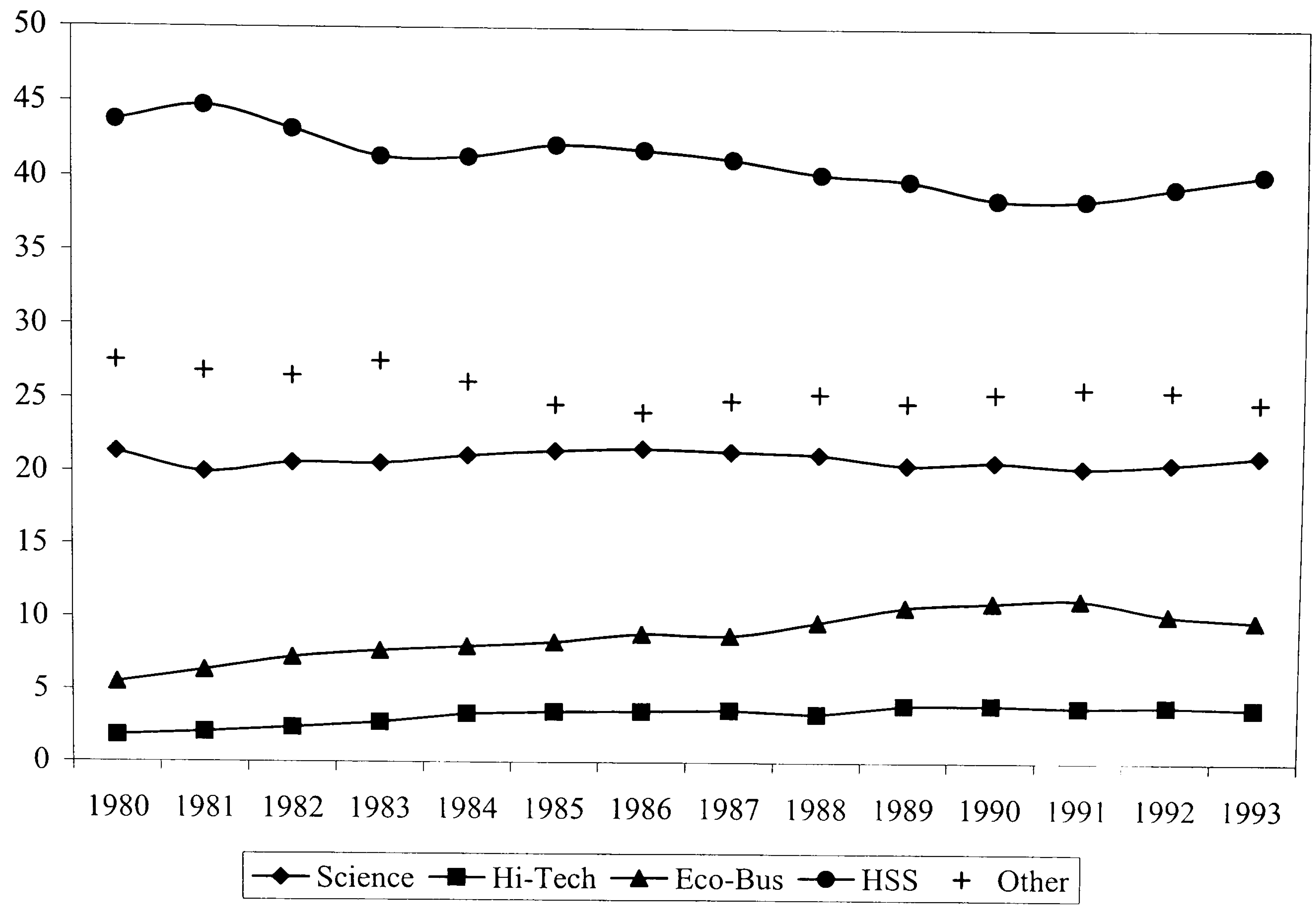


Table 5.4 Male relative earnings *premia* by degree subject and year

year		I. OLS				II. PSM-ATT				III. MNL-OLS			
		Degree subjects				Degree subjects				Degree subjects			
		Science	Hi-Tech	HSS	Other	Science	Hi-Tech	HSS	Other	Science	Hi-Tech	HSS	Other
1980	Coeff.	-0.024 **	-0.031 **	-0.034 **	0.015 **	-0.048 **	-0.050 **	-0.028 **	0.009	0.010	-0.142 **	0.064 **	0.258 **
	s.e.	-0.004	0.004	0.004	0.004	0.007	0.009	0.008	0.007	0.019	0.013	0.016	0.031
1981	Coeff.	-0.027 **	-0.021 **	-0.063 **	-0.019 **	-0.040 **	-0.046 **	-0.052 **	-0.023 *	0.149 **	0.245 **	-0.006	0.212 **
	s.e.	0.005	0.004	0.005	0.004	0.011	0.010	0.011	0.008	0.040	0.025	0.020	0.052
1982	Coeff.	-0.036 **	-0.014 **	-0.061 **	-0.001	-0.071 **	-0.037 **	-0.034 **	-0.002	-0.015	-0.150 **	-0.047 **	0.127 **
	s.e.	0.005	0.005	0.006	0.005	0.012	0.013	0.011	0.008	0.025	0.014	0.014	0.033
1983	Coeff.	-0.057 **	-0.026 **	-0.085 **	-0.008	-0.087 **	-0.061 **	-0.074 **	-0.013	-0.019	-0.137 **	-0.045 **	0.131 **
	s.e.	0.005	0.005	0.006	0.005	0.011	0.012	0.011	0.008	0.020	0.015	0.016	0.023
1984	Coeff.	-0.061 **	-0.045 **	-0.089 **	-0.028 **	-0.092 **	-0.066 **	-0.098 **	-0.056 **	-0.026	-0.167 **	-0.054 **	0.143 **
	s.e.	0.005	0.005	0.006	0.005	0.012	0.013	0.012	0.009	0.027	0.012	0.012	0.025
1985	Coeff.	-0.045 **	-0.039 **	-0.094 **	-0.027 **	-0.082 **	-0.071 **	-0.090 **	-0.049 **	0.063 **	-0.161 **	-0.049 **	0.182 **
	s.e.	0.005	0.005	0.006	0.005	0.014	0.015	0.012	0.009	0.023	0.012	0.013	0.028
1986	Coeff.	-0.080 **	-0.075 **	-0.098 **	-0.053 **	-0.124 **	-0.127 **	-0.088 **	-0.077 **	-0.057 **	-0.238 **	-0.062 **	0.145 **
	s.e.	0.006	0.006	0.007	0.006	0.015	0.019	0.016	0.011	0.021	0.013	0.013	0.021
1987	Coeff.	-0.083 **	-0.080 **	-0.097 **	-0.034 **	-0.117 **	-0.105 **	-0.099 **	-0.047 **	-0.095 **	-0.238 **	-0.054 **	0.236 **
	s.e.	0.006	0.006	0.007	0.006	0.015	0.017	0.015	0.012	0.035	0.014	0.013	0.019
1988	Coeff.	-0.069 **	-0.061 **	-0.109 **	-0.032 **	-0.100 **	-0.092 **	-0.102 **	-0.051 **	0.029	0.205 **	-0.047 **	0.196 **
	s.e.	0.006	0.006	0.007	0.006	0.015	0.014	0.015	0.011	0.024	0.013	0.013	0.030
1989	Coeff.	-0.058 **	-0.050 **	-0.100 **	0.000	-0.086 **	-0.090 **	-0.085 **	-0.016 **	-0.072 **	-0.215 **	-0.080 **	0.140 **
	s.e.	0.006	0.005	0.007	0.006	0.013	0.016	0.019	0.011	0.018	0.014	0.014	0.027
1990	Coeff.	-0.040 **	-0.041 **	-0.100 **	0.050 **	-0.101 **	-0.060 **	-0.076 **	0.025 **	0.021	-0.166 **	-0.045 **	0.224 **
	s.e.	0.006	0.006	0.007	0.006	0.014	0.017	0.015	0.011	0.027	0.019	0.019	0.030
1991	Coeff.	-0.031 **	-0.023 **	-0.098 **	0.082 **	-0.086 **	-0.076 **	-0.091	0.068 **	0.160 **	0.308 **	-0.007	0.334 **
	s.e.	0.007	0.007	0.008	0.007	0.016	0.017	0.017	0.012	0.039	0.017	0.015	0.043
1992	Coeff.	-0.050 **	-0.031 **	-0.098 **	0.059 **	-0.099 **	-0.041 **	-0.086 **	0.049 **	0.108 **	0.289 **	-0.034 **	0.300 **
	s.e.	0.007	0.007	0.008	0.007	0.016	0.022	0.019	0.013	0.039	0.020	0.016	0.044
1993	Coeff.	-0.074 **	-0.045 **	-0.106 **	0.035 **	-0.140 **	-0.092 **	-0.080 **	0.028 *	0.053 **	0.250 **	-0.057 **	0.336 **
	s.e.	0.008	0.008	0.008	0.008	0.019	0.024	0.019	0.014	0.031	0.022	0.015	0.042

** and * denote statistical significance at 1% and 5% level, respectively.

Figure 5.5 Male relative earnings*premia* by degree subject and year: OLS

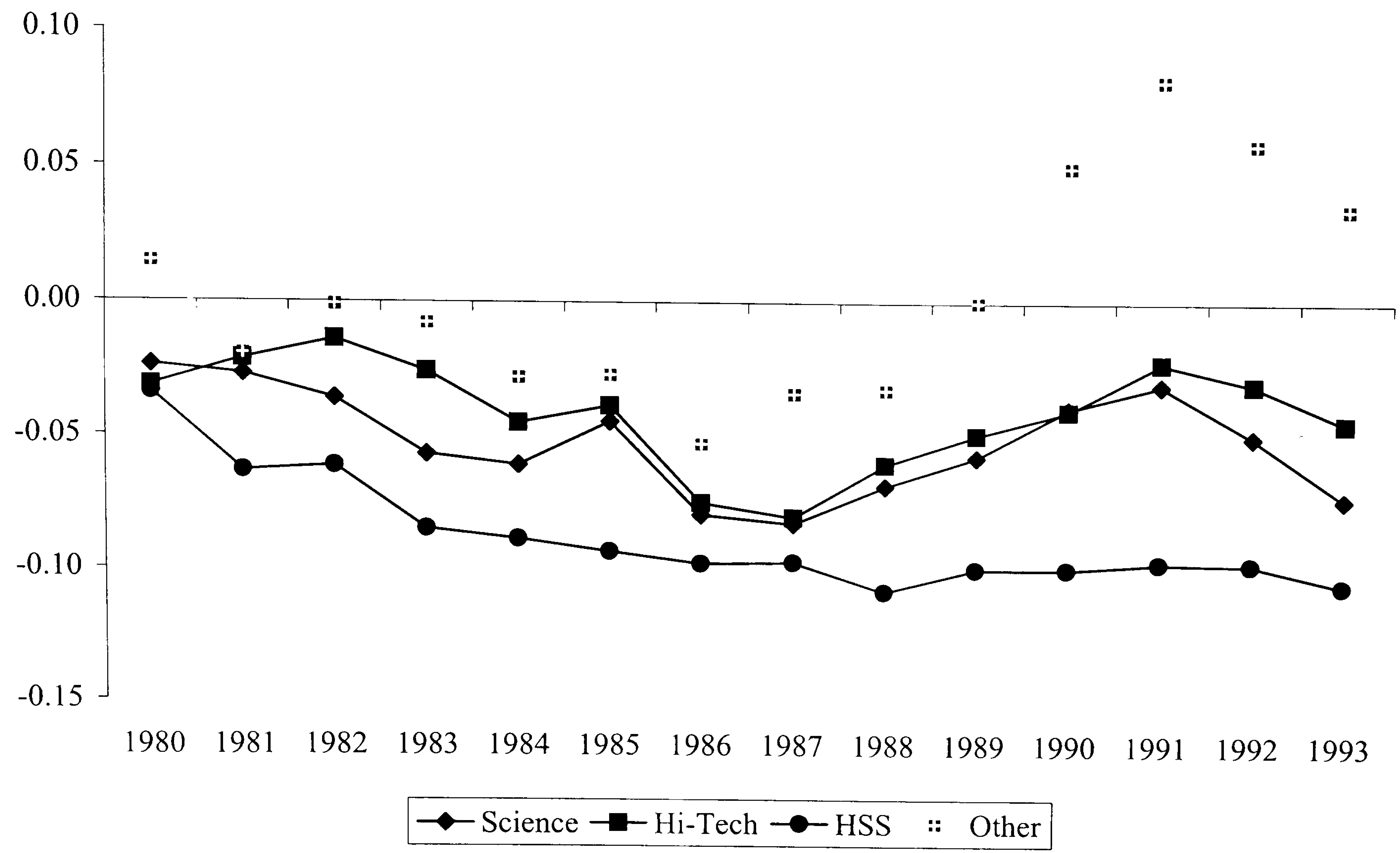


Figure 5.6 Male relative earnings*premia* by degree subject and year: PSM-ATT

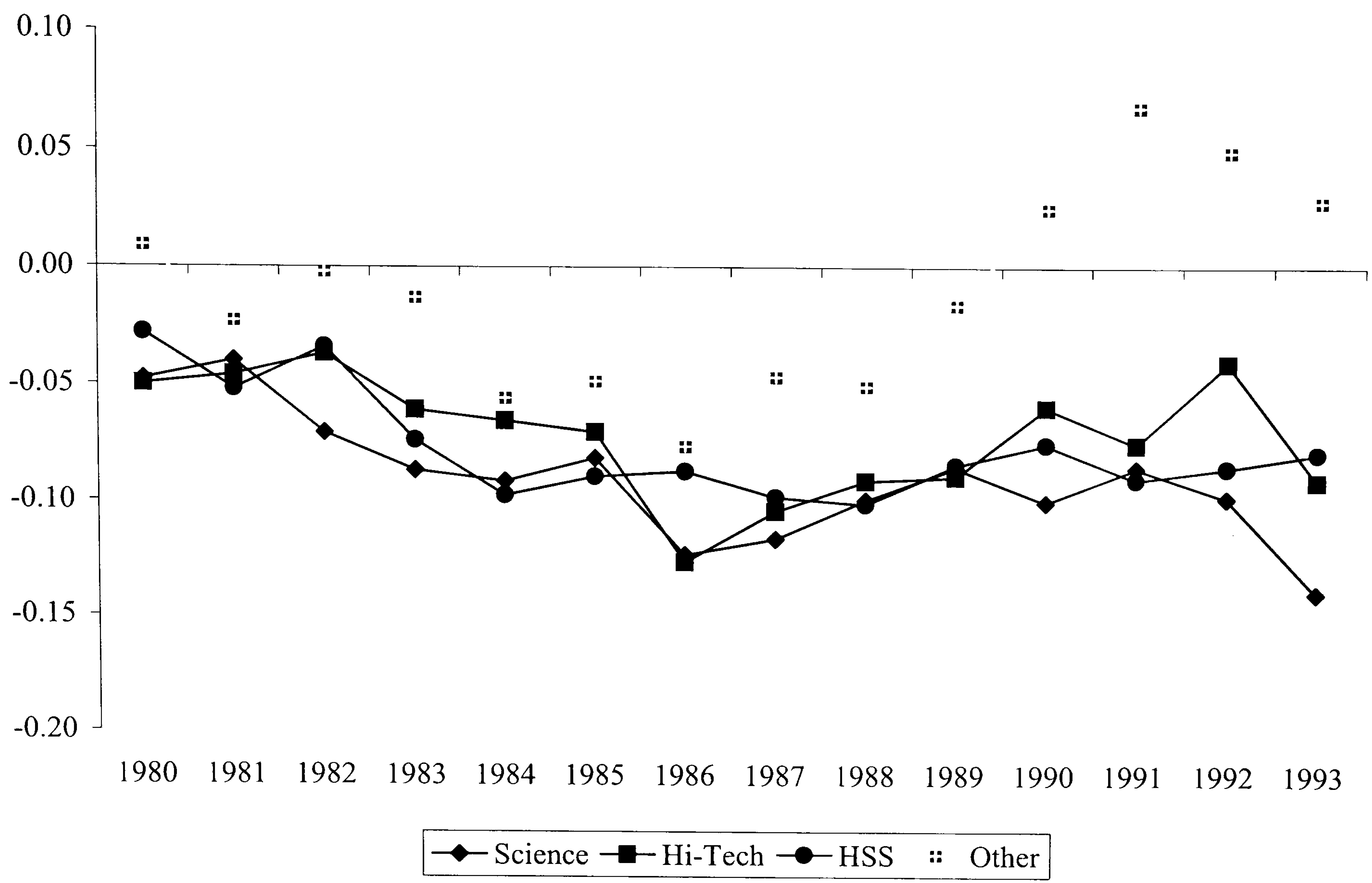


Figure 5.7 Male relative earnings*premia* by degree subject and year: MNL-OLS

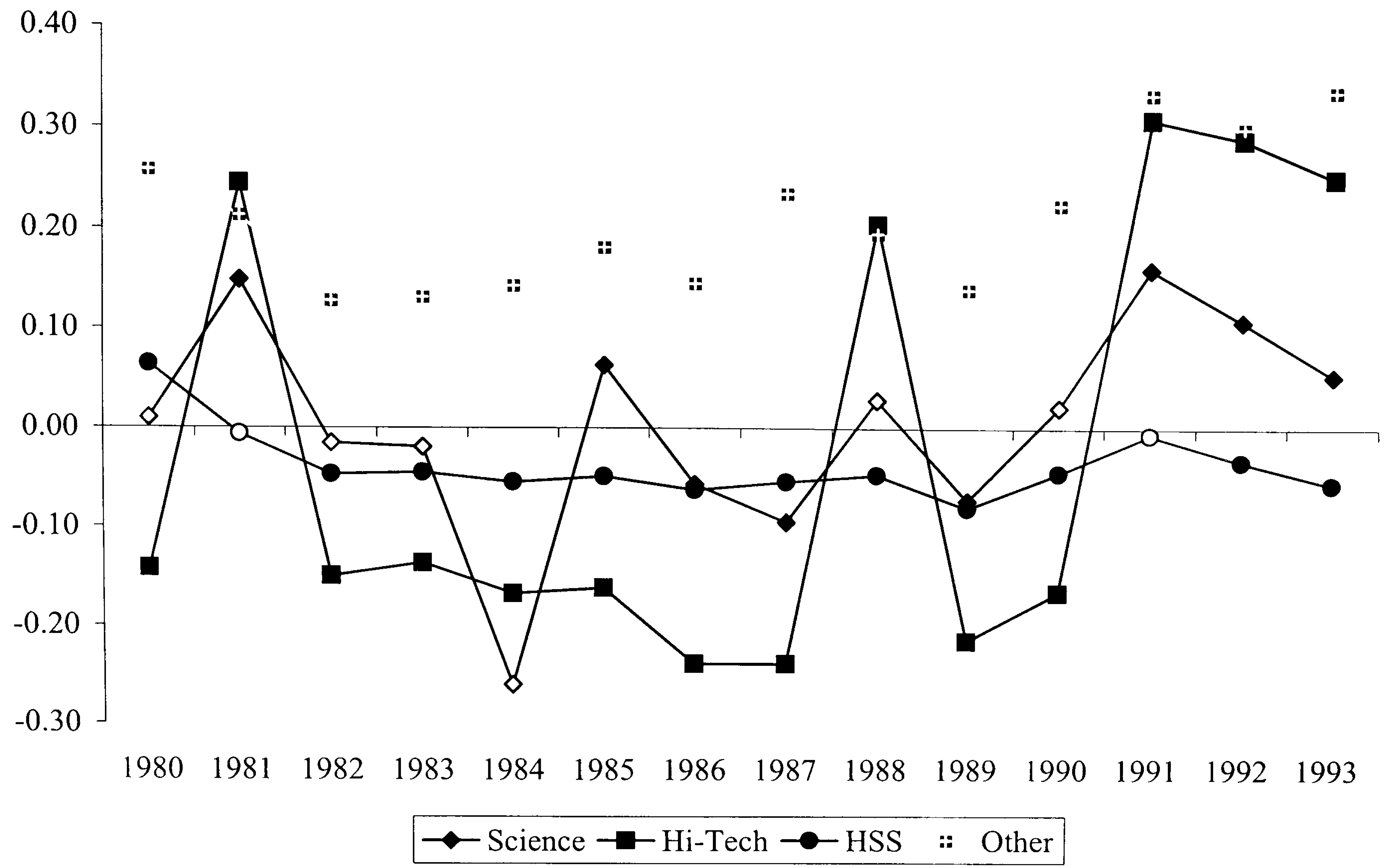


Table 5.5 Choice of identifying variables in the MNL-OLS model for males

year	ID.Vs	Earnings regression (OLS)		Subject choice (MNL)	
		LR-test (a)	p-value	LR-test (b)	p-value
1980	ALTECH, ALHSS, MATURE	$Chi^2(3)=1.27$	0.74	$Chi^2(12)=531.7$	0.00
1981	ALTECH, ALOTH	$Chi^2(2)=1.62$	0.45	$Chi^2(8)=495.2$	0.00
1982	ALTECH, ALHSS, ALOTH	$Chi^2(3)=1.48$	0.69	$Chi^2(12)=551.6$	0.00
1983	ALTECH, ALOTH	$Chi^2(2)=0.74$	0.69	$Chi^2(8)=585.7$	0.00
1984	MATURE	t=-1.22	0.22	$Chi^2(4)=48.1$	0.00
1985	ALTECH, MATURE	$Chi^2(2)=0.61$	0.74	$Chi^2(8)=422.5$	0.00
1986	ALTECH, ALHSS, ALOTH	$Chi^2(3)=1.61$	0.66	$Chi^2(12)=795.0$	0.00
1987	ALTECH	t=0.35	0.72	$Chi^2(4)=433.4$	0.00
1988	ALOTH, MATURE	$Chi^2(2)=0.95$	0.62	$Chi^2(8)=334.9$	0.00
1989	ALTECH	t=0.25	0.80	$Chi^2(4)=646.1$	0.00
1990	ALTECH, ALOTH	$Chi^2(2)=1.15$	0.56	$Chi^2(8)=758.2$	0.00
1991	ALTECH, ALOTH, MATURE	$Chi^2(3)=0.94$	0.81	$Chi^2(12)=914.3$	0.00
1992	ALTECH, ALOTH, MATURE	$Chi^2(3)=5.11$	0.16	$Chi^2(12)=1172.3$	0.00
1993	ALTECH	t=1.30	0.19	$Chi^2(4)=744.0$	0.00

(a) Likelihood ratio test for the exclusion of the ID.Vs from the earnings regressions

(b) Likelihood ratio test for the exclusion of the ID.Vs from the MNL model of subject choice

Table 5.6 Estimated coefficients of correlation (ρ)

year	Degree subjects				
	Science	Hi-Tech	Eco-Bus	HSS	Other
1980	0.257 *	0.865 *	0.368 *	-0.007	-0.691 *
1981	-0.464 *	-0.897 *	0.286 *	0.057 *	-0.694 *
1982	-0.093	0.767 *	0.051	0.021	-0.514 *
1983	-0.063	0.741 *	0.134 *	-0.012	-0.440 *
1984	-0.061	0.814 *	0.138 *	0.031	-0.569 *
1985	-0.407 *	0.812 *	0.139 *	-0.001	-0.639 *
1986	-0.062	0.832 *	0.087 *	-0.022	-0.604 *
1987	0.128	0.796 *	0.106 *	0.000	-0.748 *
1988	-0.064	-0.821 *	0.226 *	0.030	-0.490 *
1989	0.051	0.807 *	0.043	-0.001	-0.431 *
1990	-0.062	0.801 *	0.191 *	0.029	-0.374 *
1991	-0.258 *	-0.848 *	0.302 *	0.033	-0.406 *
1992	-0.213 *	-0.815 *	0.213 *	0.039	-0.417 *
1993	-0.151 *	-0.748 *	0.168 *	0.045	-0.569 *

* denotes statistical significance at 5% level

Table 5.7 Female relative earnings *premia* by degree subject and year

year		I. OLS				II. PSM-ATT				III. MNL-OLS			
		Degree subjects				Degree subjects				Degree subjects			
		Science	Hi-Tech	HSS	Other	Science	Hi-Tech	HSS	Other	Science	Hi-Tech	HSS	Other
1990	Coeff.	0.027 **	0.058 **	-0.006	-0.180 **	0.011	0.039 **	0.011	0.166 **	0.165 **	-0.141 **	0.071	0.284 **
	s.e.	0.007	0.009	0.007	0.007	0.016	0.019	0.015	0.011	0.057	0.045	0.042	0.044
1991	Coeff.	0.047 **	0.092 **	0.003	0.189 **	0.028	0.099 **	0.011	0.175 **	0.056	-0.105 *	0.039	0.020 **
	s.e.	0.007	0.010	0.007	0.007	0.019	0.019	0.016	0.012	0.055	0.054	0.050	0.005
1992	Coeff.	0.041 **	0.078 **	0.018 **	0.189 **	0.022	0.113 **	0.024	0.177 **	0.071	-0.164 **	0.062	0.223 **
	s.e.	0.008	0.011	0.008	0.008	0.019	0.026	0.017	0.014	0.060	0.067	0.066	0.094
1993	Coeff.	0.012	0.063 **	0.001	0.177 **	0.016	0.093 **	-0.004	0.186 **	-0.003	0.254 **	-0.011	0.111 **
	s.e.	0.008	0.012	0.008	0.008	0.023	0.030	0.019	0.014	0.065	0.089	0.065	0.004

* and ** denote statistical significance at 1% and 5% level, respectively.

Table 5.8 Gender differences in occupational earnings by subject and year

year	Science	Hi-tech	Eco-Bus	HSS	Other	Total
1990	0.35	0.30	0.44	0.34	0.28	0.33
1991	0.32	0.27	0.42	0.32	0.30	0.32
1992	0.29	0.26	0.40	0.29	0.26	0.29
1993	0.30	0.27	0.41	0.30	0.25	0.29

Table 5.9 Gender earnings gap between 'matched pairs' of graduates

year		Science	Hi-Tech	Eco-Bus	HSS
1990	coeff.	0.329 **	0.306 **	0.432 **	0.341 **
	s.e.	0.007	0.010	0.009	0.008
	N	7669	5251	4589	9951
1991	coeff.	0.318 **	0.274 **	0.424 **	0.316 **
	s.e.	0.008	0.012	0.018	0.009
	N	7266	4542	4523	9953
1992	coeff.	0.288 **	0.263 **	0.403 **	0.284 **
	s.e.	0.009	0.016	0.015	0.009
	N	7605	4579	4553	10754
1993	coeff.	0.293 **	0.274 **	0.403 **	0.306 **
	s.e.	0.009	0.018	0.011	0.008
	N	8313	4731	4678	11880

Table 5.10 Predicted probabilites of subject choice when male returns are assigned to females

year	Science			Hi-Tech			Eco-Bus			HSS			Other		
	M (a)	F (b)	FM (c)	M	F	FM	M	F	FM	M	F	FM	M	F	FM
1990	0.26	0.15	0.15	0.17	0.01	0.01	0.17	0.11	0.11	0.12	0.35	0.35	0.28	0.38	0.37
1991	0.24	0.15	0.15	0.16	0.01	0.01	0.17	0.11	0.11	0.14	0.35	0.35	0.29	0.38	0.37
1992	0.25	0.17	0.16	0.16	0.01	0.01	0.17	0.09	0.09	0.14	0.35	0.36	0.29	0.38	0.37
1993	0.25	0.16	0.16	0.15	0.01	0.01	0.16	0.09	0.09	0.15	0.37	0.38	0.29	0.36	0.36

(a) probabilites of subject choice predicted by the MNL-OLS model for males.

(b) probabilites of subject choice predicted by the MNL-OLS model for females.

(c) probabilites of subject choice predicted by the MNL-OLS model for females when female coefficients in the OLS equation are set equal to the corresponding male values.

Appendix 5A Variables definition

Subject studied

Science:	BIOL OBIOL CHEM PHYS OPHYS MATHS
Hi-Tech:	ENGIN COMP
Eco-Bus:	ECON BUS
HSS:	SOCIO POL OSOSCI CLAS MEUL HUM
Other:	LAW ARTS ALMED EDU COMB OTHER

Number of A-level passes by 'broad' subject area:

Mathematics (ALMATH)

Science (ALSCI): Biology, Chemistry, Physics, Other Sciences, and Statistics.

Hi-Tech (ALTECH): Computer Studies, Electronics, Mechanics, and Engineering.

Eco-Bus (ALBUS): Economics, Business.

HSS (ALHSS): English, French, German, Italian, Spanish, Other Languages, Law, Politics, Classics, Geography, and History.

Other (ALOTH) (residual group including minor subjects)

Appendix 5B The MNL-OLS model: the log-likelihood function

In this Appendix we report the expression for the log-likelihood function of our polychotomous choice model. Following Lee (1983), we estimate a simultaneous MNL-OLS model. More details about how the model is constructed can be found in Lee' paper (p. 511).

The log likelihood function for the MNL-OLS model has the following form:

$$\ln L = \sum_{i=1}^{N}\sum_{j=1}^{J}\{ S_{ij} \ln \Phi((J_j(Z_i\delta_j) + (\rho_j/\sigma_\varepsilon)(Y_{ij} - \sum_{j=1}^{J} S_{ij}\theta_j - X_i\beta))/(1-\rho_j^2)^{1/2} + \\ + S_{ij}\ln\phi((Y_{ij} - \sum_{j=1}^{J} S_{ij}\theta_j - X_i\beta)/\sigma_\varepsilon) - S_{ij}\ln\sigma_\varepsilon\}$$

(A.5.1)

where N is the number of individuals. Following Lee (1983) we assume

$$\begin{bmatrix} J_1(u_{i1}) \\ J_2(u_{i2}) \\ J_3(u_{i3}) \\ J_4(u_{i4}) \\ J_5(u_{i5}) \\ J_6(\varepsilon_i) \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \rho_1 \\ 0 & 1 & 0 & 0 & 0 & \rho_2 \\ 0 & 0 & 1 & 0 & 0 & \rho_3 \\ 0 & 0 & 0 & 1 & 0 & \rho_4 \\ 0 & 0 & 0 & 0 & 1 & \rho_5 \\ \rho_1 & \rho_2 & \rho_3 & \rho_4 & \rho_5 & 1 \end{bmatrix} \right)$$

where $J_j(u_{ij})$ for j=1,...,5 is the transformation of the non-normal stochastic term u_{ij} into standard normal, while $J_6(\varepsilon_i)$ is the transformation of $\varepsilon_i \sim N(0,\sigma^2_\varepsilon)$ into standard normal and

$$J_j(Z_i\delta_j) = \Phi^{-1}\Lambda_{ij}(Z_i\delta_j)$$

$$\Lambda_{ij}(Z_i\delta_j) = \exp(Z_i\delta_j)/(1+\sum_{j=1}^{J}\exp(Z_i\delta_j),$$

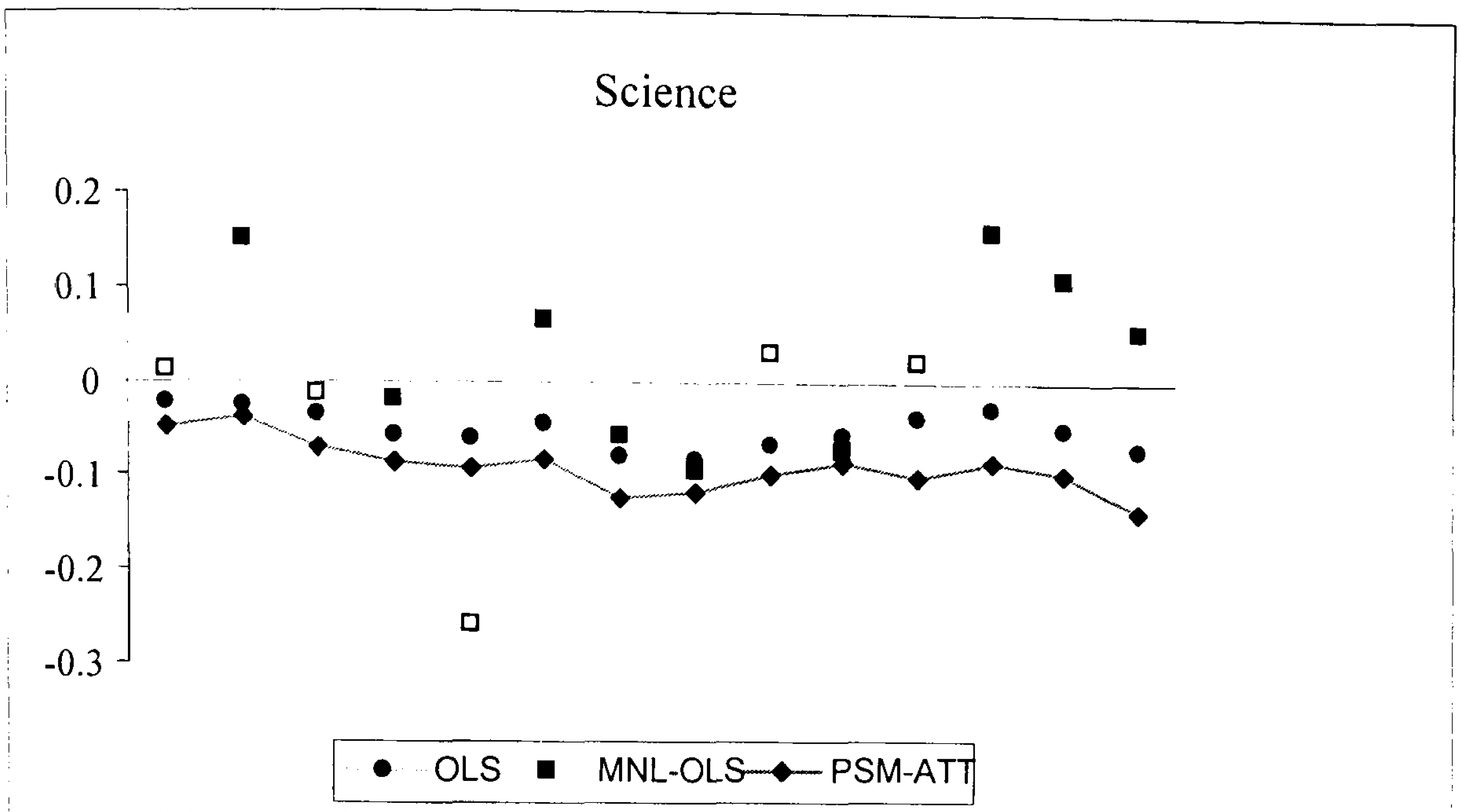
$\phi(.)$ and $\Phi(.)$ are the standard normal density and distribution functions. The main difference with respect to the original Lee's specification is that we do not estimate separate earnings regressions by subject but only one earnings regression in which the 'endogenous' degree subject dummies appear among the regressors. This specification was preferred given the specific focus of this chapter, e.g. the use of standard Mincerian log-earnings equations to estimate graduates' relative occupational earnings *premia* by subject studied correcting for the *sample selection bias* into degree course.

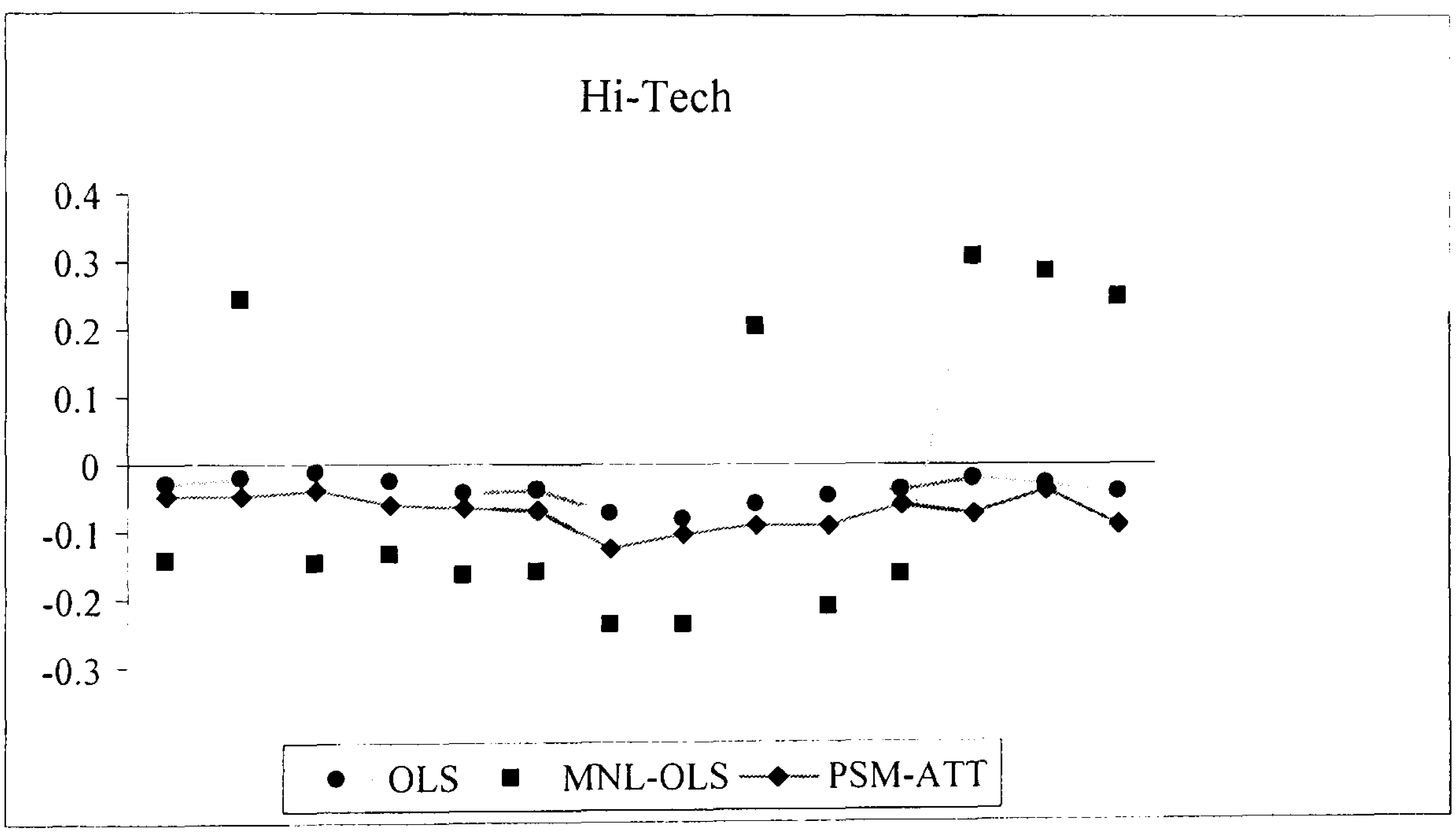
Appendix 5C Sensitivity of MNL-OLS subject *premia* to alternative identification strategies

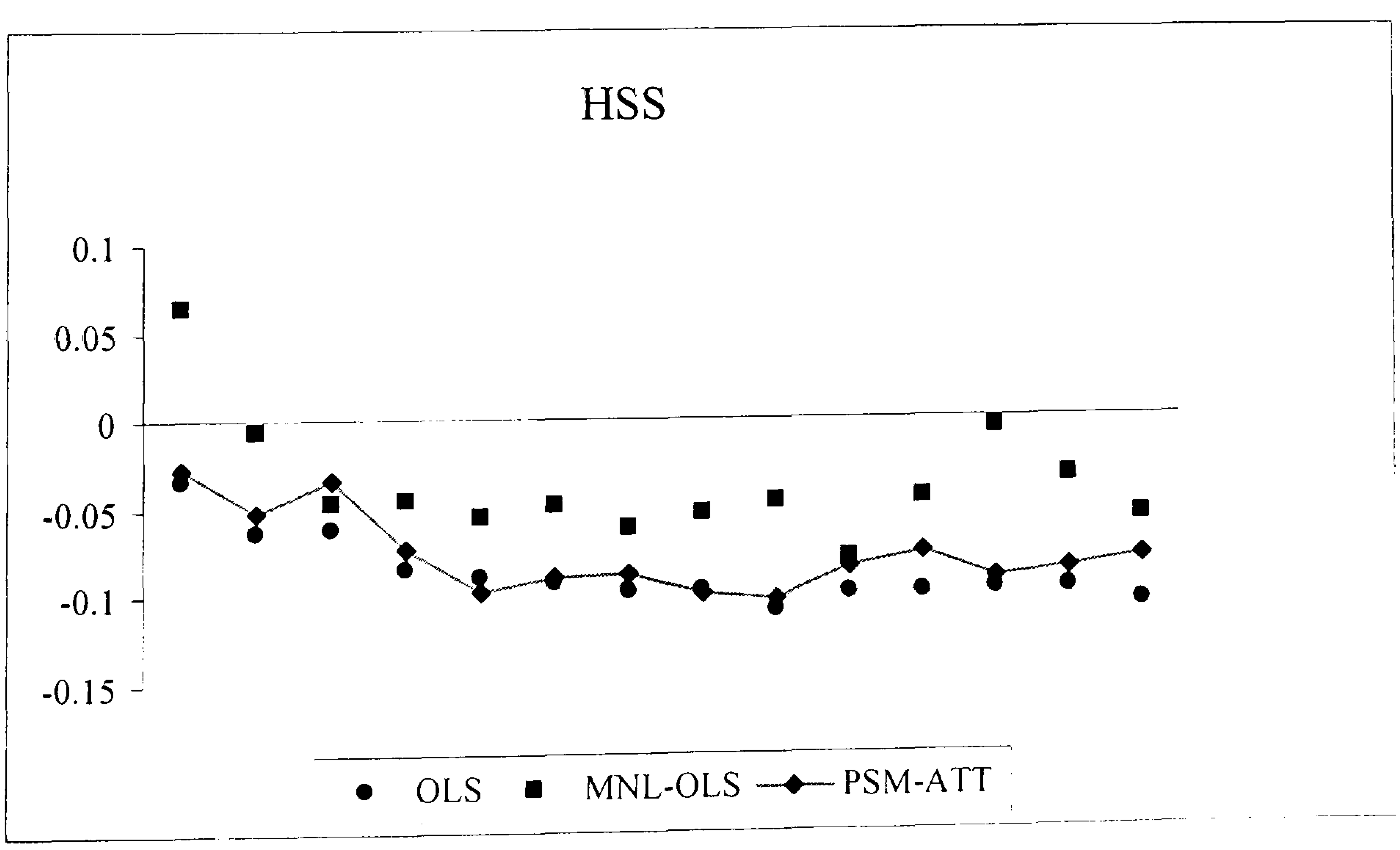
	Science		Hi-Tech		HSS		Other	
year	θ(a)	θ(b)	θ(a)	θ(b)	θ(a)	θ(b)	θ(a)	θ(b)
1980	0.009	0.010	-0.139	-0.142	0.073	0.064	0.267	0.258
1981	0.138	0.149	0.243	0.245	-0.011	-0.006	0.199	0.212
1982	-0.015	-0.015	-0.157	-0.150	-0.046	-0.047	0.131	0.127
1983	-0.019	-0.019	-0.145	-0.137	-0.043	-0.045	0.136	0.131
1984	-0.024	-0.022	-0.172	-0.172	-0.052	-0.053	0.146	0.146
1985	0.064	0.063	-0.171	-0.161	-0.050	-0.049	0.181	0.182
1986	-0.052	-0.057	-0.244	-0.238	-0.059	-0.062	0.152	0.145
1987	-0.083	-0.095	-0.248	-0.238	-0.055	-0.054	0.232	0.236
1988	-0.015	0.029	-0.216	0.205	-0.064	-0.047	0.167	0.196
1989	-0.071	-0.072	-0.230	-0.215	-0.081	-0.080	0.138	0.140
1990	0.025	0.021	-0.184	-0.166	-0.046	-0.045	0.223	0.224
1991	0.153	0.160	0.327	0.308	-0.008	-0.007	0.335	0.334
1992	0.099	0.108	0.316	0.289	-0.037	-0.034	0.299	0.300
1993	0.052	0.053	0.287	0.250	-0.058	-0.057	0.350	0.336

(a) MNL-OLS model's identification is through the functional forms only.
(b) MNL-OLS model's identification is through the functional forms and ID.Vs.

Appendix 5D Graphical comparisons between male relative earnings *premia* by subject estimated by OLS, PSM-ATT and MNL-OLS







Appendix 5E LR tests for the equality of male relative earnings *premia* by subject over time [a]

years	$Chi^2(4)$	p-value
1981	911.08	0.00
1982	584.7	0.00
1983	518.28	0.00
1984	817.79	0.00
1985	863.38	0.00
1986	1044.41	0.00
1987	919.57	0.00
1988	751.47	0.00
1989	778.84	0.00
1990	739.67	0.00
1991	833.62	0.00
1992	443.46	0.00
1993	322.79	0.00

[a] The constraints are imposed only on the coefficients of the subject dummies θ, that is:

H_0: $\theta_j(t)=\theta_j(t-1)$ j=1,.....,4

H_1: $\theta_j(t)\neq\theta_j(t-1)$

where $\theta_j(t-1)$ is taken as given. No inter-cohort equality restrictions are imposed on the coefficients of the other explanatory variables.

Chapter 6

Conclusions

The Thesis seeks to make a contribution to our current understanding of the complex relationship between higher education and the graduate labour market in the UK on both a methodological and policy level. Using a unique and comprehensive dataset on complete cohorts of individual students who left university between 1980 and 1993, the Thesis has developed along three main avenues: i) identifying the determinants of graduates' first destinations and measuring their impact on early career success; ii) comparing alternative indicators of employment-related university performance and assessing their robustness to data aggregation; iii) estimating the differences in graduates' occupational earnings by degree subject.

The first point is addressed in Chapters 2 and 3. While Chapter 2 aims to pinpoint the key factors driving the first destination outcomes of individual university students, Chapter 3 looks at how the impact of these factors has evolved over time. Given the complementarities between the analyses of the two chapters, their main conclusions are discussed jointly. The results, largely based on multinomial logistic regressions, have highlighted a number of interesting facts. First, the choice to consider an unusually broad range of possible outcomes, including the distinction between temporary and permanent employment, professional training and higher degree study, involuntary unemployment and unavailability to work, is supported by the data. Our results confirm that each of these routes reflects a unique mix of individual preferences and attributes and, therefore, they should be regarded as separate outcomes when modelling the career decisions of university graduates.

Second, there are significant gender differences in first destination outcomes especially with respect to unemployment. In 1993, male graduates were 5% more likely than females to be unemployed six months after graduation. The Oaxaca-Blinder decomposition of this differential reveals that this gap is almost entirely due to differences in unobservable factors. When looking at earlier cohorts of graduates, women's employability has been significantly less affected than men's by the economic downturn of the early 1990s. A more in-depth analysis suggests a number of concurrent explanations for the observed gender effects, including sectoral and occupational segregation, and potential discrimination in the form of affirmative–action recruitment policies favouring females.

Third, in line with previous studies (McKnight, 1999), vocational courses like Education, Allied Medicine, Computer Science, Business, and Engineering are associated with a relatively high probability of entering a permanent occupation. The probability of temporary employment is relatively high for graduates in

Sociology and Politics, while Law, Modern European Languages, Classics and Humanities graduates are more likely to enter professional training (legal and teaching, respectively). Finally, graduates in Science degrees like Chemistry, Biology, and Physics are more likely to undertake postgraduate education. Notwithstanding the persistence of a significant stratification by subject studied in the graduate labour market, employability differences between courses have generally fallen over time. It appears that these trends are related to differences in the 'quality' of employment across subjects. For instance, Engineering and Computer Science graduates who are traditionally less likely to be overqualified than other graduates are those who paid the highest price in terms of falling (rising) employment (unemployment) rates in the 1990s. This may be due either to higher career expectations or to the lack of transferable skills that make these graduates less flexible or occupationally mobile.

Fourth, individuals from a working class background (skilled manual, partly skilled and unskilled parents) are more likely to be unemployed, inactive and overqualified than graduates from professional, managerial and technical parental backgrounds. This effect is greater for males and during periods of economic slump. Interestingly, we tend to find opposite results for individuals educated at an independent school, after controlling for student social class. If one regards attendance at an independent school as an additional socio-economic factor, the joint evidence discussed above may indicate that social networks, defined as social ties to those in high-paying jobs, can represent a cost-effective recruitment strategy and a useful screening device, particularly when demand is slack. This explanation is corroborated by further evidence drawn from the subject-specific analysis. In particular, working class graduates in Law are significantly more likely to be unemployed than their wealthier peers. The influence of social and business networks may be particularly strong in legal professions, especially if one considers that a relatively high proportion (over 16%) of Law graduates have parents in legal occupations.

Fifth, degree performance has a positive and monotonic effect on an individual's probability of moving on to a higher degree course and a negative and monotonic effect on the probability to be either unemployed or overqualified. Moreover, the negative consequences on graduates' employability of getting a poor degree (lower second class honours or lower) are systematically higher during 'low phases' of the business cycle. In particular, the gradient has become much steeper during the 1990s, reflecting the fact that as degree performance has improved over the years (higher proportion of individuals graduating with at least upper second class honours) so has the stratification by academic ability in the labour market. We also find that the effect of degree performance on graduates' employability varies significantly across subjects, after controlling for university type and prior qualifications. For instance, in courses like Engineering, Computer Science, Business and Economics the employment prospects of graduates are particularly sensitive to degree class. Interestingly, these subjects also tend to be i) courses with a relatively high proportion of low achievers and ii) courses that typically offer relatively good employment prospects upon graduation. These findings suggest the existence of a positive correlation between the level of (degree class-related) risk and the size of the return (employability-related) associated with choosing a particular subject. This conclusion may have direct implications for the ongoing debate on student finance. The popularity of Engineering, Computer Science, Business and Economics degrees as courses that tend to offer good employment and pay prospects is likely to make these subjects prime candidates for 'top-up' fees (Naylor *et al.*, 2002). However, as they become more expensive, risk-averse applicants may decide to choose 'safer', but ultimately less remunerative, degrees. This need not be necessarily true if an adequate cross-subsidisation mechanism is devised to balance out the higher opportunity costs faced by less affluent students. Nevertheless, we believe that further research is required to explore the social impact of alternative funding options before market fees are introduced.

Sixth, curriculum breadth at A-level adds to the employability of Engineering and Computer Science graduates over and above the effect of degree performance and the type of university attended. These results seem to support the view that employers tend to value technical degrees more highly if graduates have some breadth of knowledge cutting across different areas of specialisation. In a period of rapid technical change, specialist skills may become more easily outdated. These findings are also informative for the current policy debate on reforming A-level curricula in the UK. Since September 2000, in response to criticism that English A-level curricula were too narrow and inadequately prepared students for the world of work, pupils are expected to take up to 5 different subjects in their first year of sixth form, while only in their last year of secondary schooling will they specialise taking the traditional three A-level subjects. Recent findings suggesting that a broader A-level curriculum does not command higher starting salaries (Dolton and Vignoles, 2001) have questioned the economic rationale of having broader curricula. Our results suggest that the absence of significant returns to a broader curriculum need not deter students from taking a larger variety of subjects at A-levels. In fact, students may perceive the extra effort made for a more rounded and differentiated curriculum worthwhile if this facilitates their transition from university to work. We also find that A-level Mathematics improves the probability of entering an occupation for Economics, Law, Humanities, Languages and Education graduates. This result supports previous evidence suggesting that A-level Mathematics positively affects future earnings (Dolton and Vignoles, 2001; Naylor *et al.*, 2002). With the exception of Economics, these subjects have typically low proportions of students taking Mathematics at A-level. Therefore, from a policy point of view, it seems important to encourage more students to take this subject, regardless of their field of study at university.

Finally, graduating from top-ranked institutions reduces the probability of early unemployment as well as the likelihood of entering 'non-graduate' occupations, particularly when the labour market is slack. Interestingly, in the early 1990s graduates from leading institutions were significantly less likely to be employed, after controlling for institutional differences in prior qualifications and degree class. These trends seem to be largely explained by the (nearly) symmetric increase in the probability of further study. The latter result may reflect the fact that leading institutions typically have a large postgraduate population. Consequently, undergraduate students are more exposed to different aspects of postgraduate education, including a greater awareness of the programmes' content and the career prospects they offer. Furthermore, when finding a job becomes more difficult because of adverse economic conditions, graduates from leading institutions may have a relatively higher 'reservation job offer'. This makes them less likely to enter 'any' occupation and more likely to stay on in higher education. In addition, cost-effective recruitment strategies may induce employers to concentrate on a smaller pool of universities with an established reputation or a record, based either on previous experience or on prejudices, for having supplied successful graduates in the past.

Employability upon graduation matters not only to individual students, but represents a pressing concern also for higher education institutions and for the central government. The sizeable share of the public budget selectively allocated to higher education every year invariably creates expectations on individual institutions to offer programmes, which are responsive to employer needs by equipping students with an effective mix of skills and tools to help them make a successful transition to the labour market. Universities are competing fiercely with each other to secure an adequate amount of resources and to retain their quality standards and international recognition. This situation has created a 'winner-take-all' syndrome at the institution level, in the sense that a growing proportion of high achievers are expected to concentrate on applying to just a few leading universities. In recent years, measuring the impact of attending a specific institution on graduates' early career outcomes and ranking universities accordingly has received increasing attention in the UK. In a typically regulated sector like education, the construction of these indicators not only serves the purpose of informing the choices of perspective students in the absence of price competition, but also represents a tool to enhance the accountability of individual institutions to the public funds received. Furthermore, it constitutes an incentive for universities to emulate best practice through peer comparison. However, the development of official indicators of employment-related university performance has proved a difficult task. The fact that the HEFCE in its annual report on performance indicators for higher education in the UK started to publish measures of employment outcomes only in 2002 about three years after these indicators were first announced by the Chancellor of the Exchequer in his 1999 Pre-Budget Report, is a testament to the difficulties encountered.

After reviewing some of the main methodological obstacles to the construction of reliable indicators examined in earlier studies (Goldstein and Spiegelhalter, 1996), Chapter 4 contributes to the literature on performance indicators by looking at the sensitivity of these indicators, and ultimately university league tables, to data aggregation. This research idea was stimulated by the evidence that university league tables ranking universities according to their success to produce employable graduates have continued to be largely based on university-level first destination information, despite individual-level data have meanwhile become available to research. First destination choices are made by individual students according to their aspirations, tastes and ability, and accounting for individual heterogeneity is important to develop reliable indicators. When student-level information is not available, a convenient but simplistic assumption, which underlies earlier studies based on university-level data, is that macro relations describe the behaviour of a fictitious 'representative' individual in the same way as the micro relations reflect the decision choices of single agents. Chapter 4 illustrates the methodology to construct alternative sets of indicators using both student-level and university-level data on 1993 university leavers respectively, based on the proportion of unemployed or inactive individuals six months after the graduation date. The resulting university rankings are then compared to assess the effects of data aggregation. Our findings reveal that a 'representative agent' approach to aggregation leads to employment-related university rankings that are significantly different from the rankings obtained from disaggregated data. We argue that these differences are the result of an aggregation bias. The validity of this conclusion is tested via a Monte Carlo experiment discussed in Section 4.7, which simulates the true data generating process of first destination decisions and then assesses the predictions of the *micro* and *macro* models against the 'true' rankings. This empirical method enables us to control for the effect of other forms of bias like misspecification. The experiment suggests that rank correlation between macro-based and true rankings does not exceed 55%. Given that universities are understandably very sensitive to the rank position they are assigned to in the tables, the evidence produced is sufficient to suggest caution about the validity of league tables based on university averages only. As a final step, the effect of using alternative aggregation procedures (Koppelman, 1975a; Train, 1986) is explored in an attempt to reduce the bias associated with the 'representative agent' approach. The results indicate that with incomplete individual-level data, aggregating over individuals after partitioning the population into more homogeneous classes can significantly reduce the representative agent bias by making some allowance at the macro level for the underlying individual heterogeneity.

Chapter 5 looks at the differences in occupational earnings by subject studied. The main contribution of this analysis to the existing empirical literature consists of comparing three alternative models that differ in the way they control for the endogeneity of subject choice. The OLS approach is used as the baseline. This first set of results is then contrasted with estimates obtained from propensity score matching methods which have become an increasingly popular technique in the evaluation literature (Rosenbaum and Rubin, 1983). Although substantially different, both methods hinge on the assumption that selection is driven solely by observable factors. Finally, we estimate a simultaneous model of graduate earnings and subject choice (Lee, 1983), which also allows for self-selection through 'unobservables'. The truly innovative aspect of this analysis is that the endogeneity of subject studied is not just acknowledged or left as a cautionary note to the reliability or interpretation of the estimated *premia* as in previous studies, but is modelled directly by estimating simultaneous equations of occupational earnings determination and subject choice. More formally, the subject effects obtained from standard log-earnings equations are corrected for student self-selection into subject groups by fitting contextually a multinomial logit regression of subject choice. This modelling strategy enables us to fully account for the correlation between the unobserved individual characteristics affecting subject choice and those affecting occupational earnings. If this correlation remains unaccounted for, subject dummies may simply pick up the effect of unobserved individual attributes rather than the 'true' earnings *premium* associated with a specific degree course. For instance, subject *premia* may incorporate the effect of idiosyncratic occupational preferences or the higher preference for non-pecuniary characteristics of the job by graduates in certain subjects.

The results suggest that earnings *premia* are in general statistically significant. Therefore, returns to a university degree estimated by standard earnings regressions controlling only for the level of educational attainment, and not for the subject studied, have to be considered only as the average return of a degree, with marked differences even across broadly defined subjects. When taking into account the potential sample selection of students into field of study based on student observable and unobservable characteristics, the dynamic patterns of the estimated earnings *premia* by degree course become considerably more volatile. In particular, we generally observe a positive selection for graduates in Economics & Business and Hi-Tech courses (except for periods of economic downturn for the latter group), while no selection effect is found for Humanities and other Social Sciences graduates. The evidence of a significant degree of correlation between the unobservable factors driving both subject choice and occupational earnings casts doubts on the reliability of rates of return estimated without accounting for the endogeneity of subject studied. In fact, earnings differences due to individual unobserved characteristics may be wrongly ascribed to the subject of degree. Moreover, earnings *premia* are likely to change over time, thus affecting the relative ranking of subjects even between consecutive years. As a consequence, studies focusing on specific cohorts of graduates may give only a very short-term account of the relative economic return to different degree subjects.

The simultaneous equation model is also used to investigate gender differences in occupational earnings and, in particular, to test the plausibility of a policy aimed at reducing the gender wage gap (Machin and Puhani, 2003) through gearing more women towards high-pay male-dominated subjects like Hi-Tech courses. Women may choose not to specialise in male-dominated fields because they receive lower returns than men from those fields (Brown and Corcoran, 1997; Paglin and Rufolo, 1990). One way to test this claim is to equalise the relative returns to subject studied and other human capital factors across genders and see if females' subject choices change accordingly. We find that the educational orientations of female graduates are unlikely to shift. Men and women's wages differ because of underlying differences in innate talent and preferences and not because of differences in academic curricula. On average, female students may decide to read relatively 'low pay' courses because they set different priorities with respect to their working life by attaching, for instance, a higher value than males to the non-pecuniary aspects of the job as opposed to high pay, or because they perceive that success in some male-dominated degrees or careers is more uncertain.

Finally, we are aware of the limitations of the Thesis. First, the USR data contain information only on students who have participated in higher education in the UK. There is no control group of non-students and, therefore, we are not able to address the issue of non-random selection into university. Consequently, all of our estimates should be interpreted as conditional on university attendance. Second, we have considered only individuals who responded to the First Destination

Survey, with the consequence that particular groups of students like individuals who failed their degree or dropped out of university, or overseas students who left the UK after graduation are largely underrepresented in our dataset. On the other hand, the main benefit of First Destination Survey data is that, despite a less than complete response rate, they do give detailed information on a much larger sample of university graduates than is available elsewhere. Third, first destination outcomes refer to arrangements that graduates made by the end of the calendar year in which they gained their degree. Career decisions made within less than six months from graduation may not necessarily reflect future career choices. However, despite this short-term horizon, we believe that first destination outcomes are still a powerful instrument to elicit longer-term career preferences and prospects, as confirmed by previous evidence (Dolton and Makepeace, 1992; Purcell and Pitcher, 1996; McKnight, 1999). Fourth, the use of occupational earnings implies the loss of any intra-occupational variation in pay, compared to individual starting salaries. However, occupational earnings have the potential advantage of being a better proxy for career earnings and, therefore, a better measure of the lifetime rate of return to a university degree compared to starting salaries.

There are a number of possible directions for future work. First, the analysis on earnings *premia* correcting for student self-selection into degree subjects could be estimated using data that report information on current rather than occupational earnings to account for intra-occupational variation in pay levels. Second, any of the analyses could be replicated for more recent cohorts of university graduates. We were unable to extend the dataset beyond 1993 because the access to the student-level archived records has been withheld since the formation of HESA in 1994-95. In particular, it would be interesting to see how first destination behaviour might have changed following the introduction of tuition fees in 1998 and the recent steps taken to reform student finance in the UK. A third extension to this Thesis is targeting postgraduates. In recent years, the rate of development of the postgraduate sector has been relatively high compared to other branches of HE in the UK (Dearing, 1997). These trends reflect the growing importance of credentials and specialist knowledge in an increasingly complex, competitive and international labour market for highly-qualified people. To date, there is limited empirical work on postgraduates based on individual-level data in the UK (Elias *et al.*, 1997; Machin and Oswald, 2000). Moreover, the analysis on undergraduates' first destinations presented in this Thesis can be used to correct for selection into postgraduate education, as individuals studying for higher degrees are likely to represent a non-random sample of the population of first degree graduates.

Bibliography

Allenby, G. and Rossi, P. (1991). "There is No Aggregation Bias: Why Macro Logit Models Work." <u>Journal of Business and Economic Statistics</u> **9**(1): 1:15

Altonji, J. G. (1993). "The Demand for and the Return to Education when Education Outcomes are Uncertain." <u>Journal of Labour Economics</u> **11**(1): 48-83.

Altonji, J. G. (1995). "The Effects of High School Curriculum on Education and Labour Market Income." <u>The Journal of Human Resources</u> **30**(3): 409-438.

Andrews, M. and Bradley, S. (1997). "Modelling the Transition from School and the Demand for Training in the United Kingdom." <u>Economica</u> **64**(255): 387-413.

Anker, R. (1997). "Theories of Occupational Segregation by Sex: An Overview." <u>International Labour Review</u> **136**(3): 315-339.

Arcidiacono, P. (2002). "Ability Sorting and the Returns to College Major." Forthcoming in <u>Journal of Econometrics</u>.

Becker, G. S. (1993). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* Chicago and London: Chicago University Press.

Belfield, C., Bullock, A., Chevalier, A., Fielding, A., Siebert, W. S. and Thomas, H. (1997). *Mapping the Careers of Highly Qualified Workers*. London: HEFCE.

Bell, D. and Elias, P. (2000). *Analysis of Pay Trends. A Teaching Profession for the 21st Century. A Report for the McCrone Inquiry.* Edinburgh: McCrone Commission.

Berger, M. C. (1988). "Predicted Future Earnings and Choice of College Major." <u>Industrial and Labour Relations Review</u> **41**(3): 418-429.

Blackaby, D. H., Murphy, P. D. and O'Leary, N. C. (1999). "Graduate Earnings in Britain: A Matter of Degree?" <u>Applied Economic Letters</u> **6**(5): 311-315.

Blinder, A. S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates." <u>Journal of Human Resources</u> **8**(4): 436-455.

Blundell, R. and Costa Dias, M. (2002). "Altenative Approaches to Evaluation in Empirical Microeconomics." CEMMAP Working Paper n. 10/02. London: Centre for Microdata Methods and Practice, University College London

Blundell, R., Lorraine, D., Goodman, A. and Reed, H. (2000). "The Returns to Higher Education in Britain: Evidence from a British Cohort." <u>The Economic Journal</u> **110**(464): F82-F89.

Blundell, R., Lorraine, D., Goodman, A. and Reed, H. (1997). *Higher Education, Employment and Earnings in Britain*. London: The Institute of Fiscal Studies.

Booth, A. (1996). *The Economics of the Trade Union*. Cambridge: Cambridge Universiy Press.

Bound, J., Jaeger, D. A. and Baker, R. M. (1995). "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable Is Weak." *Journal of the American Statistical Association* **90**(430): 443-450.

Bradley, S., Crouchley, R., Millington, J. and Taylor, J. (2000). "Testing for Quasi-Market Forces in Secondary Education." *Oxford Bulletin of Economics and Statistics* **62**(3): 357-390.

Brennan, J. L., Lyon, E. S., McGeevor, P. A. and Murray, K. (1993). *Students, Courses and Jobs. The Relationship between Higher Education and the Labour Market*. Higher Education Policy Series 21. London: Jessica Kingsley Publishers

Brown, C. and Corcoran, M. (1997). "Sex-Based Differences in School Content and the Male-Female Wage Gap." *Journal of Labour Economics* **15**(3): 431-465.

Cameron, A. C. (1990). "Aggregation in Discrete Choice Models: An llustration of Non-Linear Aggregation." in T. S. Barker and M. H. Pesaran (eds) *Disaggregation in Econometric Modelling*. London: Rutledge.

Chevalier, A. (2000). Graduate Over-Education in the UK. Discussion Paper n. 8. London: Centre for the Economics of Education.

Chevalier, A. (2002). "Just Like Daddy: the Occupational Choice of UK Graduates." Presented at the Royal Economic Society Annual Conference 2002, University of Warwick.

Chevalier, A., Conlon, G., Galindo-Rueda, F. and McNally, S. (2002). *The Returns to Higher Education Teaching*. Research Report for the Department of Education and Skills. London: Centre for the Economics of Higher Education.

Connor, H., La Valle, I., Pollard, E. and Millmore, B. (1997). *What Do Graduates Do Next?*. Brighton: Institute for Employment Studies.

Davies, S. and Guppy, N. (1997). "Fields of Study, College Selectivity, and Student Inequalities in Higher Education." *Social Forces* **75**(4): 1417-1438.

Davis, E. P. (1995). "Institutional Investors, Unstable Financial Markets and Monetary Policy." Special Paper n. 75. London: LSE Financial Markets Group.

Daymont, T. N. and Andrisani, P. J. (1984). "Job Preferences, College Major, and the Gender Gap in Earnings." *Journal of Human Resources* **19**(3): 408-428.

Dearing, R. (1997). *Higher Education in a Learning Society - Report of the National Committee of Inquiry into Higher Education - Main Report*. London: HMSO.

Deaton, A. S. and Muellbauer, J. (1980). *Economics and Consumer Behaviour*. Cambridge: Cambridge University Press.

Doeringer, P. and Piore, M. (1971). *Internal Labour Markets and Manpower Analysis*. Heath, Lexington, MA

Dolton, P. J., Greenaway, D. and Vignoles, A. (1997). "Whither Higher Education? An Economic Perspective for the Dearing Committee of Enquiry." Economic Journal **107**(442): 710-726.

Dolton, P. J. and Makepeace, G. H. (1990). "The Earnings of Economics Graduates." The Economic Journal **100**(399): 237-250.

Dolton, P. J. and Makepeace, G. H. (1992). "The Early Careers of 1980 Graduates, Work Histories, Job Tenure, Career Mobility and Occupational Choice." The Department of Employment research Paper n. 79. London: Department of Employment.

Dolton, P. J., Makepeace, G. H. and Van der Klaauw, W. (1989). "Occupational Choice and Earnings Determination: the Role of Sample Selection and Non-Pecuniary Factors." Oxford Economic Papers **41**(3): 573-594.

Dolton, P. J., O' Neill, D. and Sweetman, O. (1996). "Gender Differences in the Changing Labour Market: The Role of Legislation and Inequality in Changing the Wage Gap for Qualified Workers in the United Kingdom." Journal of Human Resources **31**(3): 549-565.

Dolton, P. J. and Vignoles, A. (1999). "The Economic Case for Reforming A Levels." Centre for Economic Performance Discussion Paper n. 422. London: CEP.

Dolton, P. J. and Vignoles, A. (2000). "The Incidence and Effect of Overeducation in the Graduate Labour Market." Economics of Education Review **19**(2):179-198.

Dolton, P. J. and Vignoles, A. (2001). "Is a Broader Curriculum Better?" Economics of Education Review **21**(5): 415-429.

Elias, P. (1999). *Moving On: Graduate Careers Three Years After Graduation: Short Report*. Manchester: CSU.

Elias P., McKnight A., Purcell K., and Wilson R. (1997). *A Study of the Labour Market for Social Science Postgraduates*. IER, University of Warwick.

Finegold, D., Keep, E., Miliband, D., Robertson, D., Sisson, K. and Ziman, J. (1992). *Higher Education: Expansion and Reform*. London: Institute for Public Policy Research.

Gallegati, M. (2002). "The Business Cycle Puzzle: Empirical Evidence from the G7." Mimeo, University of Ancona.

Goldstein, H. and Spiegelhalter, D. (1996). "League Tables and their Limitations: Statistical Issues in Comparisons of Institutional Performance." Journal of the Royal Statistical Society (Series A) **159**(3): 385-443.

Gorman, W. M. (1953). "Community Preference Fields." Econometrica **21**(1): 63-80.

Greenaway, H. and Williams, G. L. (1973). *Patterns of Change in Graduate Employment*. Research into Higher Education Monographs n. 19. London: Society for Research into Higher Education.

Greenaway, D. and Haynes, M. (2000). *Funding Universities to Meet National and International Challenges*. Nottingham: The School of Economic Policy, University of Nottingham.

Greene, W. H. (2000). *Econometric Analysis* (*4th Edition*). Prentice Hall.

Hall, S. (2002). "Financial Accelerator Effects in the UK Business Cycles." Bank of England Working Paper n. 150. London: Bank of England

Hanushek, E. A., Rivkin, S. G. and Taylor, L. L. (1996). "Aggregation and the Estimated Effects of School Resources." NBER Working paper n. 5548. Chicago: NBER

Harkness, S. and Machin, S. (1999). *Graduate Earnings in Britain,1974-1995*. Department for Education and Employment Research Report. Nottingham: Department for Education and Employment.

Harmon, C. and Walker, I. (1995). "Estimates of the Economic Return to Schooling for the United Kingdom." American Economic Review **85**(5): 1278-1286.

Harmon, C. and Walker, I. (1999). "The Marginal and Average Return to Schooling in the UK." European Economic Review **43**(4-6): 879-887.

Harmon, C. and Walker, I. (2000). "The Returns to the Quantity and Quality of Education: Evidence for Men in England and Wales." Economica **67**(265): 19-35.

Hausman, J. A. and McFadden, D. (1984). "Specification Tests for the Multinomial Logit Model." Econometrica **52**(5): 1219-1240.

HEFCE (2001). *Indicators of Employment.* Report n. 2001/21. Bristol: HEFCE

HEFCE (2002). *Performance Indicators in Higher Education.* Report n. 2002/52. Bristol: HEFCE

Hellerstein, D. (1995). "Welfare Estimation Using Aggregate and Individual-Observation Models: A Comparison Using Monte Carlo Techniques." American Journal of Agricultural Economics 77(3): 620-630.

Holzer, H. J. (1987). "Hiring Procedures in the Firm: Their Economic Determinants and Outcomes." NBER Working paper n. 2185. Chicago: NBER

Johnes, J. and McNabb, R. (2002). "Academic Standards in UK Universities: More for Less or Less for More?" Presented at the Royal Economic Society Annual Conference 2002, University of Warwick.

Johnes, J. and Taylor, J. (1990). *Performance Indicators in Higher Education.* Bristol: Open University Press.

Kelejian, H. (1995). "Aggregated Heterogeneous Dependent Data and the Logit Model: A Suggested Approach." Economic Letters **47**(3-4): 243-248.

Koppelman, F. S. (1975a). *Alternative Aggregation Procedures.* Boston, MA: Centre for Transportation Studies, MIT.

Koppelman, F. S. (1975b). *Travel Prediction with Models of Individual Choice Behaviour.* Boston, MA: Center for Transportation Studies, MIT.

Koppelman, F. S. (1976). *Guidelines for Aggregate Travel Predictions Using Disaggregate Choice Models.* Evanston, IL: The Transportation Centre, Northwestern University.

Lee, K., Pesaran, M. H. and Pierce, R. (1990). "Aggregation Bias in Labour Demand Equations for the UK Economy." in T. S. Barker and M. H. Pesaran (eds) *Disaggregation in Econometric Modelling.* London: Routledge.

Lee, L. F. (1983). "Generalized Econometric Models with Selectivity." Econometrica **51**(2): 507-512.

Leuven, E. and Sianesi, B. (2003). "PSMATCH2: STATA module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing." http://ideas.repec.org/c/boc/bocode/s432001.hml. Version 1.1.7.

Lewbel, A. (1992). "Aggregation with Log-Linear Models." Review of Economic Studies **59**: 635-642.

Lissenburgh, S. and Bryson, A. (1996). *The Returns to Graduation*. London: Department for Education and Employment.

Lynton, E. (1993). "Higher Education and Employment: The Case of Humanities and Social Sciences, a Synthesis of Countries and Expert Contributions." in D. Furth (eds) *Higher Education and Employment: The Case of Humanities and Social Sciences*. Paris: OECD.

Machin, S. and Puhani, P. A. (2003). "Subject of Degree and Gender Wage Differential: Evidence from the UK and Germany." Economic Letters **79**(3): 393-400.

Machin, S. and Oswald, A. (2000). "UK Economics and the Future Supply of Academic Economists." *Economic Journal* **110**(464): F334-49.

McFadden, D. (1973). "Conditional Logit Analysis of Qualitative Choice Behaviour." In P. Zarembka_(eds) *Frontiers in Econometrics*. New York: Academic Press.

McFadden, D. and Reid, F. (1975). "Aggregate Travel Demand Forecasting from Disaggregated Behavioral Models: the Aggregation Issue." *Transportation Research Record: Travel Behaviour and Values* **534**: 24-37.

McKnight, A. (1999). "Graduate Employability and Performance Indicators: First Destination and Beyond." In P. Elias and A. McKnight (eds) *Moving On - Graduate Careers Three Years after Graduation*. Manchester: Career Services Unit (CSU).

Micklewright, J. (1989). "Choice At Sixteen." *Economica* **56**(221): 25-39.

Mingat, A. and Eicher, J. C. (1982). "Higher Education and Employment Markets in France." *Higher Education* **11**(2): 211-220.

Montgomery, J. D. (1991). "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis." *American Economic Review* **81**(5): 1408-1418.

Murphy, A. and Shuttleworth, I. (1997). "Education, Religion, and First Destination of Recent School Leavers in Northern Ireland." *The Economic and Social Review* **28**(1): 23-41.

Nam, K. C. (1997). "A Study on the Estimation and Aggregation of Disaggregate Models of Mode Choice for Freight Transport." *Transportation Research-E* **33**(3): 223-231.

Naylor, R., Smith, J., and McKnight, A. (2002). "Why is There a Graduate Earnings Premium for Students from Independent Schools." Bulletin of Economic Research **54**(4): 315-339.

Nicholson, S. and Souleles, N. S. (2001). "Physician Income Expectations and Specialty Choice." NBER Working paper n. 8536. Chicago: NBER

Oaxaca, R. (1973). "Male-Female Wage Differentials in Urban Labour Markets." International Economic Review **14**(3): 693-709.

OECD (1993). *From Higher Education to Employment.* Paris: OECD

Paglin, M. and Rufolo, A. (1990). "Heterogeneous Human Capital, Occupational Choice, and Male-Female Earnings Differences." Journal of Labour Economics **8**(1): 123-144.

Pearson, R., Aston, J., Bates, P. and Jagger, N. (2000). *The IES Annual Graduate Review 2000: A Diverse and Fragmented Market.* IES Report n. 367. Brighton, UK: The Institute for Employment Studies.

Pissarides, C. A. (1981). "Staying-on at School in England and Wales." Economica **48**(192): 345-363.

Polachek, S. A. (1975). "Potential Biases in Measuring Male-Female Discrimination." Journal of Human Resources **10**(2): 205-229.

Preston, J. A. (1999). "Occupational Gender Segregation. Trends and Explanations." The Quarterly Review of Economic and Finance **39**(0) (*Special Issue 1999*): 611-624.

Purcell, K. and Pitcher, J. (1996). *Great Expectations: the New Diversity of Graduate Skills and Aspirations.* Manchester: Careers Service Unit (CSU).

Richards, M. and Ben-Akiva, M. (1975). *A Disaggregate Travel Demand Model.* Lexington, MA: Saxon House Studies.

Rigg, M., Elias, P., White, M. and Johnson, S. (1990). *An Overview of the Demand for Graduates.* London: H.M.S.O., Policy Studies Institute, and Institute for Employment Research.

Rochat, D. and Demeulemeester, J. L. (2001). "Rational Choice under Unequal Constraints: the Example of Belgian Higher Education." Economics of Education Review **20**(1): 15-26.

Roizen, J. and Jepson, M. (1985). *Degrees for Jobs: Employers Expectations of Higher Education.* Guildford, UK: SRHE & NFER-Nelson

Rosen, S. (1972). "Learning and Experience in the Labor Market." *The Journal of Human Resources* 7(3): 326-342.

Rosenbaum, P. R. and Rubin, D. B. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* **70**(1): 41-55.

Shuller, E. (1996). *The Future of Higher Education.* The Society for Research into Higher Education & Open University. Buckingham: SRHE/Open University Press series.

Sianesi, B. (2001). "Implementing Propensity Score Matching Estimation with STATA." Presented at the UK STATA Users Group, VII Meeting, London.

Sicherman, N. and Galor, O. (1990). "A Theory of Career Mobility." *Journal of Political Economy* **98**(1): 169-192.

Sloane, P. J., Battu, H. and Seaman, P. T. (1999). "Overeducation, Undereducation and the British Labour Market." *Applied Economics* **31**(11): 1437-1453.

Smith, B. and Campbell, J. (1978). "Aggregation Bias and The Demand for Housing." *International Economic Review* **19**(2): 495-504.

Smith, J. and Naylor, R. (2001). "Determinants of Individual Degree Performance." *Oxford Bulletin of Economics and Statistics* **63**(1): 29-60.

Smith, J., Naylor, R. and McKnight, A. (2000). "Graduate Employability: Policy and Performance in Higher Education in the UK." *The Economic Journal* **110**(464): F382-F411.

Smith, P. (1990). "The Use of Performance Indicators in the Public Sector." *Journal of the Royal Statistical Society. Series A (Statistics in Society)* **153**(1): 53-72.

Squires, G. (1990). *First Degree: the Undergraduate Curriculum.* London: Open University Press.

Stoker, T. M. (1982). "The Use of Cross-Section Data to Characterize Macro Functions." *Journal of the American Statistical Association* 77(378): 369-380.

Stoker, T. M. (1993). "Empirical Approaches to the Problem of Aggregation over Individuals." *Journal of Economic Literature* **XXXI**(December): 1827-1874.

Talvitie, A. (1973). "Aggregate Model Demand Analysis with Disaggregate Demand Models." *Proceedings: Transportation Research Forum* **14**: 583-603.

Tarsh, J. (1992). *From Higher Education to Employment (Vol. IV).* Paris: OECD

Theil, H. (1954). *Linear Aggregation of Economic Relations*. Amsterdam: North-Holland.

Train, K. (1986). *Qualitative Choice Analysis*. Boston, MA: MIT Press.

Van Daal, J. and Merkies, A. (1984). *Aggregation in Economic Research: From Individual to Macro Relations*. Dordrecht: D. Reidel Publishing Company.

Van de Werfhorst, H. G., Sullivan, A. and Cheung, S. Y. (2002). "Social Class, Ability, and Choice of Subject in Secondary and Tertiary Education in Britain." Forthcoming in British Educational Research Journal.

Van Garderen, K. J., Lee, K. and Pesaran, M. H. (2000). "Cross-Sectional Aggregation of Non-Linear Models." Journal of Econometrics **95**(2): 285-331.

Walker, I. and Zhu, Y. (2001). *The Returns to Education: Evidence from the Labour Force Surveys*. London: Department of Education and Skills.

Watson, P. L. and Westin, R. B. (1975). "Transferability of Disaggregate Mode Choice Models." Regional Science and Urban Economics **5**(2): 227-249.

Westin, R. B. (1974). "Predictions from Binary Choice Models." Journal of Econometrics **2**(1):1-16.

Whitfield, K. and Wilson, R. A. (1991). "Staying On in Full Time Education: The Educational Participation Rate of 16 Year Olds." Economica **58**(231): 391-404.

Willis, R., and Rosen, S. (1987). "Education and Self-Selection." Journal of Political Economy **85**(5): 507-536.

Wolf, A. (2002). *Does Education Matter? Myths About Education and Economic Growth*. London: Penguin Books.